

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 15:02:05 ; Search time 532 Seconds
(without alignments)
4128.393 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVBEVAGSMWAGREDIIM.....TSGPCUGLHSSQFPFPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9267905 segs, 2000278028 residues
Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlh
-Q=/abs/ABSSMB.spool/US10071838/rnatc.05042006.141923.5354/app_query.fasta.1
-DB=Published Applications NA New -QEMT=fastcap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss06h
-USER=US10071838 -CCSN 1.1 541 -@rnatc.05042006.141923.5354 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA New:

1: /SIDS5/ptodata/1/pubpna/US08 NEW PUB. seq: *
2: /SIDS5/ptodata/1/pubpna/US06 NEW PUB. seq: *
3: /SIDS5/ptodata/1/pubpna/US07 NEW PUB. seq: *
4: /SIDS5/ptodata/1/pubpna/US09 NEW PUB. seq: *
5: /SIDS5/ptodata/1/pubpna/US10 NEW PUB. seq: *
6: /SIDS5/ptodata/1/pubpna/US11 NEW PUB. seq: *
7: /SIDS5/ptodata/1/pubpna/US12 NEW PUB. seq: *
8: /SIDS5/ptodata/1/pubpna/US13 NEW PUB. seq: *
9: /SIDS5/ptodata/1/pubpna/US14 NEW PUB. seq: *
10: /SIDS5/ptodata/1/pubpna/US15 NEW PUB. seq: *
11: /SIDS5/ptodata/1/pubpna/US16 NEW PUB. seq: *
12: /SIDS5/ptodata/1/pubpna/US17 NEW PUB. seq: *
13: /SIDS5/ptodata/1/pubpna/US18 NEW PUB. seq: *
14: /SIDS5/ptodata/1/pubpna/US19 NEW PUB. seq: *
15: /SIDS5/ptodata/1/pubpna/US20 NEW PUB. seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	676	22.7	378	6	US-09-925-065A-506683
C 2	667	22.4	378	6	US-09-925-065A-506682

3	437	14.6	10490	14	US-11-011-332A-74	Sequence 74, Appl
4	380	12.7	1052	6	US-09-925-065A-16910	Sequence 16910, A
5	380	12.7	1052	6	US-09-925-065A-16911	Sequence 16911, A
6	380	12.7	1052	6	US-10-301-480-118147	Sequence 118147, A
7	380	12.7	1052	9	US-10-301-480-118148	Sequence 118148, A
8	380	12.7	1052	10	US-10-301-480-118149	Sequence 118149, A
9	380	12.7	1052	10	US-10-301-480-118150	Sequence 118150, A
10	367	12.3	1294	6	US-09-925-065A-55408	Sequence 55408, A
11	367	12.3	1294	6	US-09-925-065A-55409	Sequence 55409, A
12	367	12.3	1294	9	US-10-301-480-156647	Sequence 156647, A
13	367	12.3	1294	9	US-10-301-480-156648	Sequence 156648, A
14	367	12.3	1294	10	US-10-301-480-1770055	Sequence 1770055, A
15	367	12.3	1294	10	US-10-301-480-1770056	Sequence 1770056, A
16	348.5	11.7	1385	6	US-11-096-568A-23142	Sequence 23142, A
17	313.5	10.5	528	6	US-09-925-065A-53436	Sequence 53436, A
18	311	10.4	1820	11	US-11-096-568A-6617	Sequence 6617, A
19	297.5	10.0	804	6	US-09-925-065A-62926	Sequence 62926, A
20	297.5	10.0	804	9	US-10-301-480-164164	Sequence 164164, A
21	297.5	10.0	804	9	US-10-301-480-177573	Sequence 177573, A
22	293.5	9.8	601	9	US-10-301-480-30359	Sequence 30359, A
23	293.5	9.8	601	10	US-10-301-480-643768	Sequence 643768, A
24	292	9.8	2391	6	US-09-925-065A-69533	Sequence 69533, A
25	292	9.8	2391	6	US-09-925-065A-69535	Sequence 69535, A
26	292	9.8	2391	9	US-10-301-480-170772	Sequence 170772, A
27	292	9.8	2391	9	US-10-301-480-170773	Sequence 170773, A
28	292	9.8	2391	10	US-10-301-480-784181	Sequence 784181, A
29	292	9.8	2391	10	US-10-301-480-784182	Sequence 784182, A
30	290.5	9.7	2391	6	US-09-925-065A-69534	Sequence 69534, A
31	290.5	9.7	2391	9	US-10-301-480-170773	Sequence 170773, A
32	290.5	9.7	2391	10	US-10-301-480-784182	Sequence 784182, A
33	268	9.0	1870	14	US-11-124-368A-38	Sequence 38, Appl
34	256.5	8.6	578	14	US-11-128-061-5727	Sequence 5727, Ap
35	256.5	8.6	578	14	US-11-128-061-5727	Sequence 5727, Ap
36	256.5	8.6	578	14	US-11-128-049-2085	Sequence 2085, Ap
37	256.5	8.6	578	14	US-11-128-049-2085	Sequence 5727, Ap
38	253	8.5	2834	14	US-11-000-568-5727	Sequence 5727, Ap
39	248	8.3	3431	14	US-11-000-568-5727	Sequence 5727, Ap
40	241	8.1	10480	14	US-11-011-332A-86	Sequence 86, Appl
41	236.5	7.9	1902	9	US-10-932-182A-78197	Sequence 78197, A
42	236.5	7.9	1902	9	US-10-932-182A-78197	Sequence 78197, A
43	236	7.9	879	6	US-09-925-065A-95883	Sequence 95883, Ap
44	236	7.9	879	9	US-10-301-480-110820	Sequence 110820, A
45	236	7.9	879	10	US-10-301-480-724229	Sequence 724229, A

ALIGNMENTS

RESULT 1
US-09-925-065A-506683/c
Sequence 506683, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/269, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 506683
LENGTH: 378
TYPE: DNA

ORGANISM: Homo sapiens
US-09-925-065A-506683

Alignment Scores:
Pred. No.: 6,52e-33 Length: 378
Score: 676.00 Matches: 119
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.7% Indels: 0
DB: 6 Gaps: 0

US-10-071-838-2 (1-549) x US-09-925-065A-506683 (1-378)

Qy 361 AAlaySProgluGInGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 357 GCCAAACCCAGACAGAGGTCCTGCGCATCCAGGCTTGCGCTTCCAGTGGCGGAG 298
Qy 381 ThirleuCyblysglyAspArgGlnAlaProProglYProProAlaArgPheProArgPro 400
Db 297 ACCCTCTGCAAGGGGAGCAGGCGAGGCCCTCCAGGCCACACAGCCGGTTCGGCGGCC 238
Qy 401 IlleTPSerAlaSerProProArgAlaProArgSerSerThrProCybProglYgLYAla 420
Db 237 ATTGGTCAGCTTCCCGCCACAGGCGACCTCGTCTTCACACCTGCTCGTGGGCGCT 178
Qy 421 ValArgLysbPthrTYrProValGlyThrgInGlyValProSerProAlaLeuAlaGln 440
Db 177 GTCCGGAGACACCTACCTGTCGCGACTCAGGTCGTGCCAGGCCCGCTGGCTCAG 118
Qy 441 GlylyPProglInGlySerTPArgPheLeuGlnTPraAnsSerMetProArgLeuProThr 460
Db 117 GGAGAGCTCAGAGGTCCTGAGATTCCTGCGATGGAACTCCATGCCCGCTCCCAAG 58
Qy 461 AspleuabPValGlyGlyProTPheArgHisTYrAspPheArgGlnSerCybTP 479
Db 57 GACCTGACGTAGAGGGCCCTTGTCGCCATTATGATTCAGACAGAGCTGCTGG 1

RESULT 2

US-09-925-065A-506682/C
Sequence 506682, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925.065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 506682
LENGTH: 378
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-506682

Alignment Scores:
Pred. No.: 2.32e-32 Length: 378
Score: 667.00 Matches: 118
Percent Similarity: 99.2% Conservative: 0
Best Local Similarity: 99.2% Mismatches: 1
Query Match: 22.4% Indels: 0
DB: 6 Gaps: 0

US-10-071-838-2 (1-549) x US-09-925-065A-506682 (1-378)

Qy 361 AAlaySProgluGInGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 357 GCCAAACCCAGACAGAGGTCCTGCGCATCCAGGCTTGCGCTTCCAGTGGCGGAG 298
Qy 381 ThirleuCyblysglyAspArgGlnAlaProProglYProProAlaArgPheProArgPro 400
Db 297 ACCCTCTGCAAGGGGAGCAGGCGAGGCCCTCCAGGCCACACAGCCGGTTCGGCGGCC 238
Qy 401 IlleTPSerAlaSerProProArgAlaProArgSerSerThrProCybProglYgLYAla 420
Db 237 ATTGGTCAGCTTCCCGCCACAGGCGACCTCGTCTTCACACCTGCTCGTGGGCGCT 178
Qy 421 ValArgLysbPthrTYrProValGlyThrgInGlyValProSerProAlaLeuAlaGln 440
Db 177 GTCCGGAGACACCTACCTGTCGCGACTCAGGTCGTGCCAGGCCCGCTGGCTCAG 118
Qy 441 GlylyPProglInGlySerTPArgPheLeuGlnTPraAnsSerMetProArgLeuProThr 460
Db 117 GGAGAGCTCAGAGGTCCTGAGATTCCTGCGATGGAACTCCATGCCCGCTCCCAAG 58
Qy 461 AspleuabPValGlyGlyProTPheArgHisTYrAspPheArgGlnSerCybTP 479
Db 57 GACCTGACGTAGAGGGCCCTTGTCGCCATTATGATTCAGACAGAGCTGCTGG 1

RESULT 3

US-11-011-332A-74
Sequence 74, Application US/11011332A
Publication No. US20060024684A1
GENERAL INFORMATION:
APPLICANT: Foeckens, John
APPLICANT: Harbeck, Nadia
APPLICANT: Koenig, Thomas
APPLICANT: Maier, Sabine
APPLICANT: Martens, John
APPLICANT: Model, Fabian
APPLICANT: Nimmrich, Inko
APPLICANT: Rujan, Tamas
APPLICANT: Schmitt, Manfred
APPLICANT: Lesche, Ralf
APPLICANT: Dietrich, Dima
APPLICANT: Mueller, Volkmar
APPLICANT: Kluth, Antje
APPLICANT: Schwabe, Ina
APPLICANT: Hartmann, Oliver
APPLICANT: Adorjan, Peter
TITLE OF INVENTION: PROGNOSTIC MARKERS FOR PREDICTION OF TREATMENT RESPONSE AND/OR S
TITLE OF INVENTION: BREAST CELL PROLIFERATIVE DISORDER PATIENTS
FILE REFERENCE: 47675-99
CURRENT APPLICATION NUMBER: US/11/011.332A
CURRENT FILING DATE: 2004-12-13
PRIOR APPLICATION NUMBER: US 10/517,741
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: PCT/EP2003/010881
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: DE 10245779.4
PRIOR FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: DE 10300096.8
PRIOR FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: DE 10317955.0
PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: PCT/EP2004/014170
PRIOR FILING DATE: 2004-12-13
PRIOR APPLICATION NUMBER: EP 03090432.0
PRIOR FILING DATE: 2003-12-11
PRIOR APPLICATION NUMBER: EP 04090041.7
PRIOR FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: EP 04090380.9
PRIOR FILING DATE: 2004-09-30
PRIOR APPLICATION NUMBER: EP 04090127.4
PRIOR FILING DATE: 2004-04-01

Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 74
; LENGTH: 10490
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-011-332A-74

Alignment Scores:
Pred. No.: 4.71e-17 Length: 10490
Score: 437.00 Matches: 254
Percent Similarity: 15.4% Conservative: 4
Best Local Similarity: 15.2% Mismatches: 20
Query Match: 14.6% Indels: 1401
DB: 14 Gaps: 9

US-10-071-838-2 (1-549) x US-11-011-332A-74 (1-10490)

QY 24 LysGlyHisArgHisGlyLeuProGluAspLysGly-----ProLysProPheArgSer 41
DB 5332 CAGGGCTCCAGAGCCCAAGACCCAGATCCGGGGCGGCTCTGGAGAGCTTGGCAGCTCC 5391
QY 42 TyrAsnAspAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeuThr 61
DB 5392 GCTAACTCCAAATGCTCTATTG----- 5415
QY 62 AlaArgGluAlaLys-GlnIleArgArgGluIleSerArgLysSerLysTrpValAspHe 81
DB 5416 -----ACACCAAAATTCGGCGGAGATCAGCCGAAGAAGCAAGTGGTGATAT 5463
QY 81 tLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeu----- 94
DB 5464 GCTGGAGACTGGAGAAATATACAAAGCAGCAAAAGT-AACGTGAGAGGAGAGAGC 5522
QY 94 ----- 94
DB 5523 ACTCTCTGAGAGACAGGGGACAGGACCCATGGCTGTGGCTTGGACCATCAGCCTTC 5582
QY 94 ----- 94
DB 5583 AGAGGGTGGGCGGACACTGTCTCTCGCCAGAGACTGAGAGCTGTGCGCAGATTTC 5642
QY 94 ----- 94
DB 5643 TGCCTATTGTCGACAGCTCACCCTTGAGGAGGAAATCTGAATCTTAGGGCTGGACTAC 5702
QY 94 ----- 94
DB 5703 CCGAGCTCAAGGCTAGGGATGCGCTGTGACCTGAAGGAAGAAAGTTTCATCATCAGA 5762
QY 94 ----- 94
DB 5763 GTTTCGACTCTGAGTGTCCATCCACTTTTCAGTCTGGGAAGGAGACCTGTGCCAGC 5822
QY 94 ----- 94
DB 5823 TTGATCTCACTCTACTAGGAATCATGAGGCCAAACCGACAATTTCCAGATCCCGG 5882
QY 94 ----- 94
DB 5883 GCTGTGCTCTCAGTGGGTCAACCCGCTGGCTGTGACACAGATGTTTTTCGCCACA 5942
QY 95 ----11eapAgaIaTyTyLysGlyMetProMetAsnIleArgGlyProMetTrpSerVa 113
DB 5943 GCTCATAGATCGAGCTCAAGAGGATGCCCATGAACATCCGGGGCCCATGTGTCAGT 6002
QY 113 lleuLeuAsnIleGluLysMetLysLeuLysAsnProGlyArgTyGlnIle----- 130
DB 6003 CTTCTGAACATTGAGAAATGAAGTTGAAAAACCCGGAAGATACAGGT-ACGCTCAG 6061
QY 130 ----- 130
DB 6062 CCAAGACACAAACAGGACAGGCGCTGTGGGGGCCAGGTCTCCAGCTGAGAGGAGC 6121

QY 130 ----- 130
DB 6122 TCAAGACACCTTGGGAGCTGGGGTGAAGTCAAGATGAACCTTGGGACAGATGT 6181
QY 130 ----- 130
DB 6182 GACACAGTCAACACAGACAAACTCAGCTCTGTGACCTTCTGGCTTCACTAACAGCC 6241
QY 130 ----- 130
DB 6242 AAAATGAGCTTCTGCAGAGAAACCTTCTTGTCTCTTCTTCCGAAAGTGTGAC 6301
QY 130 ----- 130
DB 6302 TGTGGCTGACTGCACTGGGGGACGAGAGTCTTCATCTGTCTGAGACTGCTCTCC 6361
QY 131 -----MetLysGluLysGlyLysArgSerSerGluHisIleG1 143
DB 6362 TCTTGGCCCTGACCTTCAAGATCATGAGAGAGAGGAGGATCATCTGAGCAGATCCA 6421
QY 143 nArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTy 163
DB 6422 GCGCATCGACCGGGAAGTAAAGCGGACATTTAAGAGCATATATTCTTCAGGGATGATA 6481
QY 163 rGlyThrLys----- 166
DB 6482 CCGAACCAAGTAAGCTTACGGGAGCCACAGAGTCCACGAGAGATGGGTGAATGAGAG 6541
QY 166 ----- 166
DB 6542 GATGGGGCTTCCCGGAGCAGAGCCAGGGTCAACAGAGGAGATGACACAGCTGCCAA 6601
QY 166 ----- 166
DB 6602 GAGCTCTCCGGCCAGGAGAGAGCCGGACCATGAACCGAGACCTCCTGTGTTCAAG 6661
QY 166 ----- 166
DB 6662 CCTTGGCCAGACTGGAACATGTGGGCGCAGAACCCAGAGAGATCTTGAGAGATGAAG 6721
QY 166 ----- 166
DB 6722 GCAACAACAAATATCATGCAATGTGAAGGTCTCTCTGACCCATGGGACCAT 6781
QY 166 ----- 166
DB 6782 GGTAGGACCCAGGGAGGGTGCAGAGATGAGGGGCCATGAGGCCCCCCCAGGCAACAGT 6841
QY 166 ----- 166
DB 6842 GACAGACCAAAATGCTGGAGAAATTAAGGGTCTCTGAAACTCTCATCCAGATCCGTGG 6901
QY 166 ----- 166
DB 6902 AACATGACATGCAACAGCAGCTTGGCAGCCGTTGGGAGTGGCTCACAAAGCTGATG 6961
QY 166 ----- 166
DB 6962 GACTTGAACACACATCCCAAGTGTCAAGATATTGAACCACTGATTTGCAAACTGA 7021
QY 166 ----- 166
DB 7022 CATCAATGAAGAACAGCATGCGAGGTTCATGCTTGACTCTCTGCTCACTCACACGGA 7081
QY 166 ----- 166
DB 7082 GCTTGGGAGCGGCTTCAACACGGGGATGGGAGAGAGAGGCTGTCTCTTCAAA 7141
QY 166 ----- 166
DB 7142 CGAAGACCCAGTGAAGAAAGGGAACGAGCGGTGATGCGGCAAGAACGTGGGTGATC 7201

QY	166	-----	166
Db	7202	CTAGATGCATTTTGCTGAGGGA	CAGAGCCAGACCCAAATAAGCTACAGATAGATTCC 7261
QY	166	-----	166
Db	7262	CATTCCTAGGCCATTCTGGAAAAAGCCAAACCA	CAGGGACTGAGAGCAAGTCTGGGTGCC 7321
QY	166	-----	166
Db	7322	CAGGGGCTGACGGATCGGGGAGAGGCTGGGTG	CATAGGGGCCACCCTGGAAGCTTGAGG 7381
QY	166	-----	166
Db	7382	ATGAAGAGTCCGCCAGAGGGGCTGGAGCGGTG	CCGGAGACTTCGACATTGTTT 7441
QY	166	-----	166
Db	7442	GGAACCGTGGAGGAACCTGTACCCACAGACTG	AACTGGCCGTGTGCMAACTGAAAAA 7501
QY	166	-----	166
Db	7502	AAAAAAAAAATCATTAGAGTGAAAAGSATCAGG	CAAGTCACTGTACACTGGCTATT 7561
QY	166	-----	166
Db	7562	TGCATGTCAAGATGTGATTTTACTGAAACAT	TTCTTCAGAGTCTGAGGCCCTGAAGA 7621
QY	166	-----	166
Db	7622	GCTCACTGCTTATCTGTGTAACAATCTGAAC	CTGAATGGGATTTGCTGTTAGGCTTTGT 7681
QY	166	-----	166
Db	7682	AGACAAAGTGAATTAAACAACATCTGCACAAA	CAAAACCAAGCCCCCTTCTGTCTTC 7741
QY	167	----GlnArgGluLeuLeuHisIleLeuLeuAla	TyrGluGluTyrAsnPro----- 182
Db	7742	CTAGCAGCGGGGAACCTACTCCACATCTCTG	CAATATGAGAGATACAAACCCGGTGAAGTA 7801
QY	182	-----	182
Db	7802	TTCCCGGAGTGAGGTTCCCGGGCCATATTTC	ATATTGACAGAGTGGGTGTCTGTGG 7861
QY	182	-----	182
Db	7862	GGGTGTCCTGTTGCTTCTTTAAAGTTAGTAT	TTGTGTACCCACGAGATATAGAGGTAGGA 7921
QY	182	-----	182
Db	7922	TGTCAAGCTCACCGCTGGCATTAACCTCCAA	GAGAGGGGTGTCTCAAGGGGTCAAGCTG 7981
QY	182	-----	182
Db	7982	AGACACAGAGAGTCAGGGCTTGAACTCTGTG	TGTCACTGGGCTGTACCACTACTCTC 8041
QY	182	-----	182
Db	8042	AGAAACAAGAAATGACGCCCTCTCTCTGGGG	CTGCCCAAGCCCAAGGACTTGGCAGCAT 8101
QY	182	-----	182
Db	8102	CGCACACAGATGTGTCTATTCAGACAGACAT	TTTTTGGCAAGGTGTGAAGTGCCTGATGGA 8161
QY	182	-----	182
Db	8162	CTTGGCTCTTGTCAATGAATGAATGTGATCT	CTGAGAAAGCTCTTTTTCAGAGAGAGCC 8221
QY	182	-----	182
Db	8222	TCTCTTCAGAGGAAGCTCTTCAAGTCAACT	CTTGCCCTCTCCAAATGACATGAGTCTTCCC 8281
QY	183	-----	184

[illegible]


```
Db 9361 TGGCTAGTCCAGCCAGGCGCTGACCTGGGAGCTCGGGTTCTCCATGAGCTGGAGTTG 9420
Qy 252 ----- 252
Db 9421 GTTTCCTTCTCTGAGAGAGACAGAGGACAGGAGTGGGGCCCACTCCCGAGA 9480
Qy 252 ----- 252
Db 9481 GCAGGGCAAGGAGGTGTGTCCACCGGAGTGTGGGAAGTGAAGTGTGTGGGAGC 9540
Qy 252 ----- 252
Db 9541 TCTGACACCGCCAGTGTCTGCACTAGGGGAAGGCTTTCAGAGGCCCTGAAAGAGG 9600
Qy 253 ----- 253
Db 9601 AGGTTTTTAAAGGACGCCAGTGGCTGAGACCTCTGTGCTTCCATCAGAGCAAGAAAG 9660
Qy 258 spleuCyglYInCySerProleuGlyCySleuIleArgIleLeuIleAsp-Gly- 276
Db 9661 ATCTATGTGGGCAAGTGTTCCTGAGGCTGCTGCTCATTCGGAATTTGATGACGGGCTAA 9720
Qy 276 ----- 276
Db 9721 GGAGCATAGGAGACCCCTGGCTCAGGGACCTTCTTGCCTGCAATGCCCTGCTTCCC 9780
Qy 276 ----- 276
Db 9781 AGCCCGGGGTCTGGCTCACTCCAGCCACAGAGAGCTGAGCGGGTCCCAAGAGACA 9840
Qy 276 ----- 276
Db 9841 CACAAGCAAAACCCTCTGCCCAAGGGGGTCAATCCAGGGCCATGGCTGGGGCTCAGGCC 9900
Qy 276 ----- 276
Db 9901 CAGCCTCATGGGAGACTGGGCGAGACCCGACTTGAGAGGGCTCAGGGAAAGCTCAAGC 9960
Qy 276 ----- 276
Db 9961 CCTGGGCAAGCCCTCTCTCCAGAGGACCATCCCACTCAATGATGCCCCCATGAG 10020
Qy 276 ----- 276
Db 10021 GAGCTTCAAGACCTTGTCTGACCCAGCGTCTTGAGAGGCTCAGGCCACCTCATGGGAA 10080
Qy 276 ----- 276
Db 10081 GGTCACTGACTGAGACTGAAGCCCAAGTGTGGAGCTGAGGCCACAGGCCAGCC 10140
Qy 277 ----- 277
Db 10141 TGGAGAGACAGGTTCTTTCACACCTGCTGTCCCAAGATCTCTCGGGCTCACCTTG 10200
Qy 284 ArgLeuTTPAspValTyrLeuValGluGluAlaLeuMetProIleThrArgIle 303
Db 10201 CGCTGTGGACCTGTATCTGGTGAAGGCAAGCGCTTGATGCCGATTAACAAGATC 10260
Qy 304 AlaphelysValGlnGlnIlyArgLeuThrIlyeThrSerArg 317
Db 10261 GCCTTTAAGGTTCAAGCAGAGTAACTGCTACGTGTGCCAGCGG 10302

RESULT 4
US-09-925-065A-16910
; Sequence 16910, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
```

```

; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,946
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16910
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-16910

Alignment Scores:
Pred. No.: 2,05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 6 Gaps: 13

US-10-071-838-2 (1-549) x US-09-925-065A-16910 (1-1052)
Qy 277 IleSerLeuGlyLeuThrLeuArgLeuTTPAspValTyrLeuValGluGlyGlnAla 296
Db 7 ATCTCTCGGGGTCAACCTGCGGCTGTGGAGCTGTATCTGTAGAAGCGAAACAGCGG 66
Qy 297 LeuMetProIleThrArgIleAlaPheIlyValGln- 309
Db 67 TTGATGCCGATTAACAAGAAATCGCCTTTAAGTTCAAGAGTAACTACGTGTGCCAG 126
Qy 309 ----- 309
Db 127 CGGGGCTGGGAGACCCCTGGGCTCAGACCCCGACTGAGGCCGAGGAGCACTTCTCACT 186
Qy 309 ----- 309
Db 187 GTTCCTATGATGCTCKTGTCTGTGGCCAGAGGAGTCTGGCCAGTGGGCTGGGCAAGAC 246
Qy 309 ----- 309
Db 247 ACTGTACACAGAGCCATCCCAATGACCATGACCATGAAAGTGAAGTGTGTGAGCA 306
Qy 309 ----- 309
Db 307 CTTCCCTGTCCGAGTGCCTCCCGAGCAGACGATCTCTGTGTATCTGACACTGGGGT 366
Qy 309 ----- 309
Db 367 GGCACAAAAGATCCGGACCGCCCAAGTGAAGACTGAAGTGGCCACCGGATATGAGCT 426
Qy 310 ----- 310
Db 427 GTGACCATTCACAGTAATCTCCCTGCTGATATCCACCTGTCCCTAGAGGGCTCAC 486
Qy 313 rlyeThrSerArgCyglYProTTPAlaArgPheCybaaATGPhaValAspThrTTPAl 333
Db 487 GAAGAGCTCAGGTGTGGCCCGTGGCAAGCTTTTGCAACCGGTGTGTGATACCTGGGC 546
Qy 333 aArgAspGluAspThrValLeuIlyeHisLeuArgAlaSerMetIlyeLeuThrArgIly 353
Db 547 CAGGAGTGAAGACATGTCTCAAGCATTTAGGGCTCTTATGAAGAACTTAACAAGAA 606
Qy 353 sGlnGlyAspLeuProProPro-AlaIlyeProGluGlnIlyeSerIlyeAspArgPro 373
Db 607 GCAAGGGGAGCTGCCACCCCGAGGTG-----GGCTTCAGTGCATGTCCCTC- 653
Qy 373 AlProIleSerArgIlyeIlyeThrLeuCybelysAspArgGlnAlaPro-ProGly 392
```

```
Db 654 --CCCATGTGACCCCTTGGGGTAGTCACTAGTAGGGA-----GTGCCCCGA 699
Qy 393 ProProAlaArgPheProArgProIleTyrSerAlaSerProProAlaProArgSer 412
Db 700 CCCCAACCCCTACTACTGCGGCTT-----TCTCTTCACTTTTCTTCTT----- 744
Qy 413 SerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnGly 432
Db 745 -----CCTCTTCTCTCTGGAAGT-CTAAGAAAGTACAGAGAGCCGAGTCTCAGAGG 797
Qy 433 ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTyrArgPheLeuGlnTyr 452
Db 798 AGGGCTCAGTGCCT-----GTATACTGG 821
Qy 453 AsnSerMetProAlaGlyLeuProThrAspLeuAspValGluGlyProTyrPheArgHisTyr 472
Db 822 ACATGCTGTGCACCC----- 836
Qy 473 AspPheArgGlnSerCysTyrValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db 837 -----AGAGGGGAGATGTGGCAGA----- 857
Qy 493 TrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db 858 -----CCCTCAACAGAGCCCTCCACTTTCACAGGTGTCTCCCTCTCC 902
Qy 513 SerArgGlnGly-----ThrProPheArgAlaArgAspGlu 524
Db 903 CCCTCGAGGGCCCTCCAGATTACTAGACGAGCCAGACCATTTGTGGAGAGAGCCCGCC 962
Qy 525 GlnProCys--AlaProThrSer-----GlyProCysLeuGlyGlyLeuHis 539
Db 963 CCTCCCTGCAAGCAGCCACAGCCTTCAGAGAGCAGAGAGGCCCTCACTC-----CTGC 1016
Qy 539 IsLeuGlnSerSerGlnPheProPro 547
Db 1017 ACGCCT-----CCTCA 1027

RESULT 5
US-09-925-065A-16911
; Sequence 16911, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16911
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-16911

Alignment Scores:
Pred. No.: 2,05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
```

```
Query Match: 12.7% Indels: 210
DB: 6 Gaps: 13
US-10-071-838-2 (1-549) x US-09-925-065A-16911 (1-1052)
Qy 277 IIsSerLeuGlyLeuThrLeuAlaGlyLeuTyrAspValTyrLeuValGlyGlnAla 296
Db 7 ATCTCTCGGGCTCACCCCTGCGCTGTGGAGAGTGTATCTGGTAGAAGGAGCAAGCGC 66
Qy 297 LeuMetProIleThrArgIleAlaPheLeuValGlnGln----- 309
Db 67 TTGATGCCGATTAACAAAGATGCTTTAAGGTTCAGAGATTAAGTCTACGTGTGCCAG 126
Qy 309 ----- 309
Db 127 CGGGGCTGGGAGACCTGGGGGTCAAGACCCGACGTGGCCGAGGAGCACTTCTCAGACT 186
Qy 309 ----- 309
Db 187 GTCTCATGATCTCTGTGTGGCCAGAGGAGGTGCGCAGTGGGCTGGCAGAGAC 246
Qy 309 ----- 309
Db 247 ACTGTGACACGAGCCCATCCCAATGACCCAGATGAAGTGCAGAGTGTGTAGCA 306
Qy 309 ----- 309
Db 307 CTTCCTGTCCGGATGCGCCCCAGACCAAGTCTCTGTGTATATCTGACACCTGGGGT 366
Qy 309 ----- 309
Db 367 GGCCACAAAGATCCGGACCGCCAGTAGAGACTGAAGTGCCACGGGATATGAGCT 426
Qy 310 -----Lys-ArgLeuThr 313
Db 427 GTGACATTCACAGTAATCTCCCTGCGCTGATATCCACCTGTCCCTAAGAGGCTCAC 486
Qy 313 rLySerThrSerArgCysGlyProTyrPheAlaArgPheCysAsnArgPheValAspThrTyrAl 333
Db 487 GAGAGCTCCAGGCTGTGGCCCGTGGGACACTTTTTCACACGGTGTCTGATACCTGGGC 546
Qy 333 aArgAspGluAspThrValLeuValLeuValArgAlaSerMetLeuValLeuThrArgLy 353
Db 547 CAGGATGAGAGACACTGTCTCAAGCATTTAGGGCTCTATGAAGAACTTAACAAAGAAA 606
Qy 353 gGlnGlyAspLeuProProPro--AlaLysProGlnGlnGlySerSerAlaSerArgPro 373
Db 607 GCAGGGGAGACCTGCACCCCAAGTG-----GGCTCCAGTGCATGTCTCCCT- 653
Qy 373 aLProAlaSerArgGlyGlyLeuThrLeuGlyGlyAspArgGlnAlaPro--ProGly 392
Db 654 --CCCATGTGACCCCTGCGGGTAGTCACTAGTAGGGA-----GTGCCCCGA 699
Qy 393 ProProAlaArgPheProArgProIleTyrSerAlaSerProProAlaProArgSer 412
Db 700 CCCCAACCCCTACTACTGCGGCTT-----TCTCTTCACTTTTCTTCTT----- 744
Qy 413 SerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnGly 432
Db 745 -----CCTCTTCTCTCTGGAAGT-CTAAGAAAGTACAGAGAGCCGAGTCTCAGAGG 797
Qy 433 ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTyrArgPheLeuGlnTyr 452
Db 798 AGGGCTCAGTGCCT-----GTATACTGG 821
Qy 453 AsnSerMetProAlaGlyLeuProThrAspLeuAspValGluGlyProTyrPheArgHisTyr 472
Db 822 ACATGCTGTGCACCC----- 836
Qy 473 AspPheArgGlnSerCysTyrValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db 837 -----AGAGGGGAGATGTGGCAGA----- 857
```

QY 493 TrpGlnAlaGluHisProAlaGluArgSerAlaPheAlaAlaProSerThrAsp 512
DB 858 -----CCTCCAAACAAGCCCCCTCCACTTCCACGGTGTCTCCCTCC 902
QY 513 SerAspGlnGly-----ThrProPheArgAlaArgAspGlu 524
DB 903 CCTCCAGAGGCGCTCCAAAGTACTAGACGAGCCAGACCATTTGTGGAGACCCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCysLeuCyseGlyLeuH 539
DB 963 CCTCCCTGCAGAGACCCACAGCCTCAGAGACGACAGAGCCCTCACTC-----CTGC 1016
QY 539 IseuGluSerSerGlnPheProPro 547
DB 1017 ACGCT-----CCTCCA 1027

RESULT 6
US-10-301-480-118147
; Sequence 118147, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: PaasSEQ for Windows Version 4.0
; SEQ ID NO 118147
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-118147

Alignment Scores:
Pred. No.: 2,05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 9 Gaps: 13

US-10-071-838-2 (1-549) x US-10-301-480-118147 (1-1052)

QY 277 IleserLeuGlyLeuThrLeuArgLeuTPAspValTyrLeuValGluGlyGluAla 296
DB 7 ACTCTCTCGGGGCTCACCTCGCGCTGTGGAGGTATCTGTAGAAAGCGAAGCGG 66
QY 297 LeuMetProIleThrArgIleAlaPheLeuValGlnGln----- 309
DB 67 TTGATGCCATTAACAAGATCGCTTTAAGTTACACAGACAGTAAGTCTAGTGTGCCAG 126
QY 309 ----- 309
DB 127 CGGGGCTGTGGAGCCCTGGGGTACAGACCCGAGCTGGCCGAGGAGCATTCTCTCACACT 186
QY 309 ----- 309
DB 187 GTCTCATGATCCKTCTTCTGTGGCCAGAGGAGGTGTGGCCAGGTGGGCTGGGAGAGC 246
QY 309 ----- 309
DB 247 ACTGTGACACGAGCCCATCCCCACATGACCAACCAAGTGAAGTGCAGAGTGTGTAGACA 306
QY 309 ----- 309
DB 307 CTTCCTGTCCGAGTCCGCCAGCCCGACAGCAGTCTCTGTGTATATCTGACACCTGGGGT 366

QY 309 ----- 309
DB 367 GGCACAAAAGATCCGGACCCGCCAGTGAAGACTGAAGTGGCCACGGGAATGAGCT 426
QY 310 -----Lys-ArgLeuTh 313
DB 427 GTGACCATTCACAGTAACCTCCCTGGGCTGATATCCACCTGTCTCCCTAGAGGGCTCAC 486
QY 313 fLyvThrSerArgCyseGlyProTTPAlaArgPheCyseAenArgPheValAspThrTPAl 333
DB 487 GAAGAGGTCCAGGTGTGGCCGTGGGACAGCTTTTTCACACCGGTTGCTTATATCCTGGGC 546
QY 333 aArgAspGluAspThrValLeuLeuYsHsleuArgAlaSerMetLeuYsleuThrArgGly 353
DB 547 CAGGAGTGAAGACACTGTGCTCAAGCATCTTAAGGCTCTATAGAAGAACTAACAGAAA 606
QY 353 sgInGlyAspLeuProProPro--AlaYsProGlnGlnGlySerSerAlaSerArgProV 373
DB 607 GCAAGGGGAGCTGCCACCCCAAGTG-----GGCTCCAGTGCATGTCCCT- 653
QY 373 alProAlaSerArgGlyGlyLyvThrLeuCyseGlyAspArgGlnAlaPro-ProGly 392
DB 654 --CCCATGTCAACCTCTGGGGTGTAGTCAGTAGTGGGA-----GTGCCCGGGA 659
QY 393 ProProAlaArgPheProArgProIleTTPSerAlaSerProProArgAlaProArgSer 412
DB 700 CCGGCAACCTTACTACTGGGCT-----TCTCTTCACCTTTCTTCT- 744
QY 413 SerThrProCyseProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnGly 432
DB 745 -----CCTCTTCTCTCTGACGCTAAGAAATACAGAGGCCACCGGCTCTCAGGGC 797
QY 433 ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTPArgPheLeuGlnTP 452
DB 798 AGCGGCTCAGTGGCT-----GTATACTGG 821
QY 453 AenSerMetProArgLeuProThrAspLeuAspValGluGlyProTTPPheArgHsTyr 472
DB 822 ACATGCTGTGCAGCG----- 836
QY 473 AspArgGlnSerCyseTPValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
DB 837 -----AGAGGGGAGATGTGGCAAGA----- 857
QY 493 TrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
DB 858 -----CCTTCCAAAGCCCTCCCACTTCCACGAGTGTCTCCCTCC 902
QY 513 SerAspGlnGly-----ThrProPheArgAlaArgAspGlu 524
DB 903 CCTCGCAGAGGCGCTCCAAAGTACTAGACGAGCCAGACCCATTTGTGGAGACCCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCysLeuCyseGlyLeuH 539
DB 963 CCTCCCTGCAGACCCACAGCCTCAGAGACGACAGAGCCCTCACTC-----CTGC 1016
QY 539 IseuGluSerSerGlnPheProPro 547
DB 1017 ACGCT-----CCTCCA 1027

RESULT 7
US-10-301-480-118148
; Sequence 118148, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598

PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ. ID NOS: 1226818
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 118148
LENGTH: 1052
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-118148

Alignment Scores:
Pred. No.: 2,05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: Gaps: 13

US-10-071-838-2 (1-549) x US-10-301-480-118148 (1-1052)

```
Qy 277 lleserleuglyleuthrleuargleutrpapvaltyrleuvalgluglinala 296
Db 7 ATCTCTCGGGCTCACCTCGCTGCGAGCGTATCTGTGAAGAGCGAAGCGGCG 66
Qy 297 leuacproliethrargilealaphelyvalglnlin 309
Db 67 TTGATGCCGATACAGAAATCGCTTTAAGTTACAGAGATAGTCTACGTGCCAG 126
Qy 309 ----- 309
Db 127 CGGGGCTGGGAGCCCTGGGGTGACAGCCCGACTGCGCCGAGGAGACTTCTACACT 186
Qy 309 ----- 309
Db 187 GTCTCATGATCTCTGTTCTGGCCAGAGAGGTCTGCGAGGTGGCTGGGACAGAC 246
Qy 309 ----- 309
Db 247 ACTGTGACACCGAGCCATCCGCCACATGACCCAGATGAAGTGAGAGTGTGGTGA 306
Qy 309 ----- 309
Db 307 CTTCCTGTCCGATCGCCGCCAGCCAGACAGTCTCTGTATATCTGACACCTGGAGT 366
Qy 309 ----- 309
Db 367 GGGCAAAAGATCGGCGCCGCGGAGAGAGTGAAGTGGCCAGGAGATGAGCT 426
Qy 310 ----- 310
Db 427 GTGACATATCCAGGTAATCTCCCTGCGCTGATATCCACCTGTCCTAGAGCGCTGAC 486
Qy 313 rlyethrserargcygilypTOTPALargPheCyabanaargPhevalaapthTTPA 333
Db 487 GAAGACGTCCAGGTGTGCGCCGCGGAGCGTATTTTGGCAACCGGTTCGTGATACCTGGGC 546
Qy 333 aArgAspGluapPheThrValleuValleuArglaasethelyslyleuThrArgly 353
Db 547 CAGGATGTAGACACTGTGTCAAGCACTTAAGGCTCTATGAAGAACTAACAGAGAA 606
Qy 353 sglnglylAapleupProPPro-alaIysProgluglInglySerSerAlaSerArgPro 373
Db 607 GCAAGGGGAGCTGCGCAACCCCAAGTG-----GGCTCCAGTCCCATGTCCTCT 653
Qy 373 alProAlaSerArglylylylythrleuCyalyseglyAapArgGlnAlaPro-Progly 392
Db 654 --CCCATGTCACTCTGGGGTACGTAGTACGAGGGA-----GTGCCCGGGA 699
Qy 393 ProProAlaArgPheProArgProIletrPserAlaSerProProArgAlaProArgSer 412
Db 700 CCCCAGAACCTACTACCTGGGCT-----TCTCTTCACTTTTCTTCTCT----- 744
Qy 413 SerThrProCybProglylyAlaValaArgGluAapThrTyProValglyThrGlnly 432
```

```
Db 745 -----CCTCTCTCTCTGAGCT-CTAAGAAAGTACAGAGGCCACCGGCTCTCAGGGC 797
Qy 433 ValProSerProAlaValaGlnGlnGlyProGlnGlySerThrArgPheleuGlnTrp 452
Db 798 AGGCGCTCAAGTGGCT-----GTAATCTGG 821
Qy 453 AamSerwerProArgleuProThrAapleuAapValgluglyProtrPheArgHlaTyR 472
Db 822 ACATGCTGTGCACCC----- 836
Qy 473 AapPheArgGlnSerCystrPValArgAlaIleserGlnIuAapGlnleuAlaProCyb 492
Db 837 -----AGAGGGGAGTGTGGCAAGA----- 857
Qy 493 TrpGlnAlaIuHlSerProAlaGluArgValaArgSerAlaPheAlaIaProSerThrAsp 512
Db 858 -----CCTCCAAAGCCCTCCCACTTCCAGGCTGTCTCTCTCC 902
Qy 513 SerAapGlnly-----ThrProPheArgAlaArgAapGlu 524
Db 903 CCTTCGAGGGCCCTCAACTACTAGACAGCCAGACCATTTGTGGAGACCCCGCC 962
Qy 525 GlnProCyb--AlaProThrSer-----GlyProCybLeuCyalylyleuH 539
Db 963 CCTCCCTGCAAGACCCACAGCTCAGAGAGCAGAGAGGCCCTCACTC-----CTGC 1016
Qy 539 lAeugluserSerGlnPheProPro 547
Db 1017 ACGCT-----CTTCA 1027
```

RESULT 8

US-10-301-480-731556 Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of single Nucleotide Polymorphisms
FILE REFERENCE: 108827,137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
NUMBER OF SEQ. ID NOS: 1226818
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 731556
LENGTH: 1052
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-731556

Alignment Scores:
Pred. No.: 2.05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: Gaps: 13

US-10-071-838-2 (1-549) x US-10-301-480-731556 (1-1052)

```
Qy 277 lleserleuglyleuthrleuargleutrpapvaltyrleuvalgluglinala 296
Db 7 ATCTCTCGGGCTCACCTCGCTGCGAGCGTATCTGTGAAGAGCGAAGCGGCG 66
Qy 297 leuacproliethrargilealaphelyvalglnlin 309
Db 67 TTGATGCCGATACAGAAATCGCTTTAAGTTACAGAGATAGTCTACGTGCCAG 126
Qy 309 ----- 309
```

Db 127 CGGGGCTGGGAGCCCTGGGGTCAACCCCGACTGGCCGAGGGCAGCTTCTCACACT 186
 Qy 309 ----- 309
 Db 187 GTCTCATGATCCKCTGTCTGTGGCCAGAGGAGGTCTGCGCAGGTGGCTGGCAGAC 246
 Qy 309 ----- 309
 Db 247 ACTGTGACACGAGCCCATCCCCCACTAGACCCAGATGAATGAATGAGAGTGTGTGACCA 306
 Qy 309 ----- 309
 Db 307 CTTCCTGTCCGATGCCCCCAGCAGACTCTCTGTGTATATCTGACACTGGGGT 366
 Qy 309 ----- 309
 Db 367 GGGCAGAAAAGATCCGGACCCGCCAGTAGAGACTGAATGCGCCAGGATATGAGCT 426
 Qy 310 -----Lys-ArgLeuTh 313
 Db 427 GTGACCATTCGAGTAATCTCCCTGGCTGATATCCACCTGTCCCTAGAGCGCTCAC 486
 Qy 313 rlysthrSerArgCy6glyProTTPAlaArgPheCyAsnArgPheValAspThrTTPAl 333
 Db 487 GAAGAGTCCAGGTGGCCCGTGGGACGTTTTTGCAACCGGTTCTGTATACCTGGGC 546
 Qy 333 aaArgapGluAspThrValLeuLysHileuAlaArgPheValAspThrArgly 353
 Db 547 CAGGAGTGAAGACACTGTGTCTCAAGCATCTTAGGGCTCTATAGAAATCAACAGAAA 606
 Qy 353 sgInglYAspLeuProProPro-AlaLysProGluInglYSerSerAlaSerArgProV 373
 Db 607 GCGGGGGAGCTCCACCCCGAGCTG-----GGCTCAGGCGCATGTCCCT- 653
 Qy 373 alProAlaSerArglylylyThrLeuCyelysglYAspArgInlaPro-ProGly 392
 Db 654 -CCCATGTCAACCTCGGGGTAGTCAAGTAGAGGA-----GTCCCGGGA 699
 Qy 393 ProProAlaArgPheProArgProIleTPSerAlaSerProProAlaProArgSer 412
 Db 700 CCCGCAACCTTACTACTGGGCT-----TCTCTTCACTCTTTCTTCTCT 744
 Qy 413 SerThrProCySerProGlylylyValaValArgGluAspThrTyProValGlyThrInglY 432
 Db 745 -----CTCTTCTCTCTGACT-CTAAGAAATGACAGAGGCCCGCTCTCAGGGC 797
 Qy 433 ValProSerProAlaLeuAlaInglYlyProGluInglYSerTPArgPheLeuGlnTP 452
 Db 798 AGGCGCTCAGTGGT-----GTAATCTGG 821
 Qy 453 AasnSerMetProArgLeuProThrAspLeuAspValGluInglYProTPArgPheArgHsTYr 472
 Db 822 ACATGCTGTGACGC----- 836
 Qy 473 AspPheArgGlnSerCystrpValArgAlaIleSerGlnLysPglInleuAlaProCyS 492
 Db 837 -----AGAGGGGGAGTGGGCGAGA----- 857
 Qy 493 TTPGlnAlaGluHileProAlaGluValArgSerAlaPheAlaAlaProSerThrArg 512
 Db 858 -----CCTTCACAAAGCCCTCCCACTTCCACGAGTCTCTCTCTCC 902
 Qy 513 SerAspGlnGly-----ThProPheArgAlaArgAspGln 524
 Db 903 CCTTCGACAGGGCTCTCAAGTTACTAGACGAGCCAGACCATTTGTGGAGACCCCGCC 962
 Qy 525 GlnProCyS--AlaProThrSer-----GlyProCySLeuCySglyLeuH 539
 Db 963 CCTTCCTGACAGACCCAGCTCTCAAGAGACAGACAGAGCCCTCCTC-----CTGC 1016
 Qy 539 IseuGlnSerSerGlnPheProPro 547

Db 1017 ACGCT-----CTTCA 1027
 RESULT 9
 US-10-301-480-731557
 ; Sequence 731557, Application US/10301480
 ; Publication No. US20060057564A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
 ; FILE REFERENCE: 108827.137
 ; CURRENT APPLICATION NUMBER: US/10/301,480
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 122618
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 731557
 ; LENGTH: 1052
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-301-480-731557
 Alignment Scores:
 Pred. No.: 2.05e-14 Length: 1052
 Score: 380.00 Matches: 133
 Percent Similarity: 34.7% Conservative: 10
 Best Local Similarity: 32.3% Mismatches: 60
 Query Match: 12.7% Indels: 210
 DB: 10 Gaps: 13
 US-10-071-838-2 (1-549) x US-10-301-480-731557 (1-1052)
 Qy 277 ILeSerLeuGlyLeuThrLeuArgLeuTPAspValTYrLeuValGluInglYInla 296
 Db 7 ATCTCTCTGGGCTACCTCGCTGGCTGTGGACGTATCTGTAGAAAGCGAAGCGG 66
 Qy 297 LeuMetProIleThrArgIleAlaPheLysValGlnGln----- 309
 Db 67 TTGATGCCATTAACAAGAAATCGCTTTAAGTTCAAGCAGTAAGTCTAGTGTGCCAG 126
 Qy 309 ----- 309
 Db 127 CGGGGCTGGGAGACCCCTGGGGTCAAGCCCGACTGGCCGAGGAGCTTCTCACACT 186
 Qy 309 ----- 309
 Db 187 GTCTCATGATCCTCTGTCTGTGGCCAGAGGAGGTCTGAGCAGGTGGCGCAGAC 246
 Qy 309 ----- 309
 Db 247 ACTGTGACACGAGCCCATCCCCCACTAGACCCAGATGAATGAATGAGAGTGTGTGACCA 306
 Qy 309 ----- 309
 Db 307 CTTCCTGTCCGATGCCCCCAGCAGACTCTCTGTGTATCTGACACTGGGGT 366
 Qy 309 ----- 309
 Db 367 GGGCAGAAAAGATCCGGACCCGCCAGTAGAGACTGAATGCGCCAGGATATGAGCT 426
 Qy 310 -----Lys-ArgLeuTh 313
 Db 427 GTGACCATTCGAGTAATCTCCCTGGCTGATATCCACCTGTCCCTAGAGCGCTCAC 486
 Qy 313 rlysthrSerArgCy6glyProTTPAlaArgPheCyAsnArgPheValAspThrTTPAl 333
 Db 487 GAAGAGTCCAGGTGGCCCGTGGGACGTTTTTGCAACCGGTTCTGTATACCTGGGC 546
 Qy 333 aaArgapGluAspThrValLeuLysHileuAlaArgPheValAspThrArgly 353

Db 547 CAGGATGAGGACATGCTGCTCAAGCATCTTAGGGCTCTATGAAGAACTTAACAAGAA 606
Qy 353 sGInGlyAspLeuProPro-ALaLysProGInGInGlySerSerLAserArgProv 373
Db 607 GCAGGGGAGCTGCTGACCCCAAGGTG-----GGCTCCAGTGCATGTCCCTC 653
Qy 373 aLProLAserArgGlyLysThrLeuCyLysGlyAspArgLAlaPro-ProGly 392
Db 654 --CCCATGTCACTCTGGGGTACTCAGTACTAGGGA-----GTGCCCGGA 699
Qy 393 ProProLAlaArgPheProArgProLLeTTrSerLAserProProArgLAlaProArgSer 412
Db 700 CCGCAACCCCTACCTACCTGGGCT-----TCCTCTCACTTTCTCTCTC----- 744
Qy 413 SerThrProCyAspProGlyGlyAlaValArgLAspThrTyrrProValGlyThGInGly 432
Db 745 -----CCTCTTCTCTCTGAGCT-CTAAGAAAGTACAGAGAGCCACCGGTCTCAGGGC 797
Qy 433 ValProSerProAlaLeuAlaGInGlyGlyProGInGlySerTrpArgPheLeuGlnTrp 452
Db 798 AGCGCTCAAGTGCCT-----GTAATCTGG 821
Qy 453 AsnSerMetProArgLeuProThrAspLeuAspValGInGlyProTrpPheArgHisTrp 472
Db 822 ACATGCTGTGACGC----- 836
Qy 473 AspPheArgGlnSerCyTrpValArgLAlaLAserGInGInGInGInGInGInGInGInGIn 492
Db 837 -----AGGAGGGGATGTGGCAAGA----- 857
Qy 493 TrpGInLAlaGInHisProLAlaGInArgValArgSerLAlaPheLAlaLAserThrAsp 512
Db 858 -----CCTCCACAGAGCCCTCCCACTTCCACAGGTCTCTCTCTCC 902
Qy 513 SerAspGInGly-----ThrProPheArgLAlaArgAspGln 524
Db 903 CCGTCGAGGGCCCTCCCAATTACTAGACAGAGCCAGCCATTTGTGGAGAGCCCGCC 962
Qy 525 GlnProCyAs-ALaProThSer-----GlyProCyLeuCyAsGlyLeuH 539
Db 963 CCGTCCTGCAAGCAGCCAGCCCTCAGAGAGAGAGAGAGCCCTCACTC-----CTGC 1016
Qy 539 IsLeuGlnSerSerGlnPheProPro 547
Db 1017 ACGCT-----CCTCCA 1027

RESULT 10
US-09-925-065A-55408
; Sequence 55408, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55408
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-925-065A-55408
Alignment Scores:
Pred. No.: 1.53e-13 Length: 1294
Score: 367.00 Matches: 143
Percent Similarity: 32.9% Conservative: 22
Best Local Similarity: 28.5% Mismatches: 86
Query Match: 12.3% Indels: 252
DB: 6 Gaps: 16

US-10-071-838-2 (1-549) x US-09-925-065A-55408 (1-1294)
Qy 202 ProGInGInuAspLAlaPheTrpAlaLeuValGInLeuLeuLAserGInuArgHisLeu 221
Db 34 CAGGCTCATGGGAGAGCTGGGCGAGAGAGCTGAGAGGCTGAG-----GGAAAGCTC 90
Qy 222 GInGlyPheHisSerProAsnGlyGlyThValGInGlyLeuGInArgGInGInuHis 241
Db 91 AAGCCCTGGGCAAGCCCTCT-----CTCCAGAGCCACATCCCTAC 132
Qy 242 -----ValValAlaThrSerGlnPro----- 248
Db 133 TCAATGATGTCCTCCATAGAGAGCTTCAAGACCTTGTCTGACCCAGCTCTGAGGG 192
Qy 249 -----LysThrMetGlyHisGInAspLysLysAspLeuCyAs 260
Db 193 CTCAGGAGACCCCTCATGGGAGAGAGTCACTGACTCAGAGACTG-AAAGCCCACTGT 251
Qy 261 GInGInCySerProLeuGly-----CyLeuLeuArgLLeuLLe 274
Db 252 GCGAGCTGAGAGCCAGAGCCAGCCCTGAGAGAGAGAGTCTTTCACACTGTCTGCC 311
Qy 275 AspGlyLLeuSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrrLeuValGInGly 294
Db 312 CCACAGATCTCTCTCGAGCTCACCTCGCTGAGAGAGCTGATCTGTAGAAAGGAGAA 371
Qy 295 GInAlaLeuMetProLLeuTrpArgLAlaPheLysValGln-Gln----- 309
Db 372 CAGGCTGTGATGCGGATACAAAGATCGCTTTAAGGTTCAGTAGAGTAAGTACAGTGT 431
Qy 309 ----- 309
Db 432 GCCAGAGGGGCTGTGGAGAGCCCTGGGGTCAAGCCCGACTGGCCGAGGGAGACTTCT 491
Qy 309 ----- 309
Db 492 CACACTGTCTCATGATCTCTGTCTGTGGCCAGAGGAGAGTCTGGCAGGTGGCTGG 551
Qy 309 ----- 309
Db 552 CAGGACATGAGACAGGAGCCATTCGCCACATGACCCAGATGAAGTGAAGTGTGG 611
Qy 309 ----- 309
Db 612 TGAGCACTTCCTGTCCAGATCAACCCCGAGCCAGACTTCCTGTATATCGAGCGC 671
Qy 309 ----- 309
Db 672 TGGGTGGCCACAAAGAGATCCGAGCCGCCCAAGTGGAGACTGAAGTGGCCAGGGGTA 731
Qy 310 -----LysAs 311
Db 732 TGAGCTGTGACATTCAGAGTAATCCCTGTGGCTGATATCAACCTGTCTCTAAGAGC 791
Qy 311 GLeuThrLysThrSerArgCyAsGlyProTrpAlaArgPheCyAsAspArgPheValAsp 331
Db 792 CCGTCAGAGACGTCACAGGTGTGCGCCGTGGGCAAGTTTGGAAACGGTTCTGTGATGC 851
Qy 331 rTrpAlaArgAspGlnuAspThrValLeuLysHisLeuAspLAserMetLysLysLeuTh 351
Db 852 CTGGGCGAGGATATGACACTGTGCTCAAGCATCTTAAGGCTCTTAAGAAACTAAC 911
Qy 351 rArgLysGInGlyAspLeuProPro-ALaLysProGInGInGlySerSerLAserA 371

```

Db      912 AAGAAACAGGGGAGCTGCACCCAGGATG-----GGCTCCAGTGCATGT 959
Qy      371 rgrProValProAlaSerArgGlyIlyThrLeuCyblyGlyAspArgGlnAlaPro- 390
Db      960 CCCCT---CCCATGTCACTCTGGGGTACTCAATAGTACGGGA-----GTGC 1004
Qy      391 ProGlyProProAlaArgPheProArgProIleThrSerAlaSerProProArgAlaPro 410
Db      1005 CCGGACCCGCAACCTACTACTGGGCT-----TCCTTTCACTTTTCTTCT 1055
Qy      411 ArgSerSerThrProCybProGlyIlyAlaValArgGlnAspThrTyProValGlyThr 430
Db      1056 -----CCTCTTCTCTCTGACT-CTAAGAAATACAGAGGCCACCGGTCT 1102
Qy      431 GlnGlyValProSerProAlaLeuAlaGlnGlyIlyProGlnIlySerThrPhePheLeu 450
Db      1103 CAGGGACGGGCGCTCAAGTGCCT-----GTA 1126
Qy      451 GlnThrAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProThrPheArg 470
Db      1127 TACTGCACATGCTGTGCACG----- 1147
Qy      471 HisTyraSerPheArgGlnSerCybTrpValArgAlaIleSerGlnIleAspGlnLeuAla 490
Db      1148 -----AGAGGGGAGATGTGGCGAGA----- 1168
Qy      491 ProCybTrpGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAlaAlaProSer 510
Db      1169 ----- 1174
Qy      511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCybAlaProThr 530
Db      1175 -----AACAGCCCCCT----- 1186
Qy      531 SerGlyProCybLeuCybGlyLeuHisLeuGlnSerSerGlnPheProGlyPhe 549
Db      1187 -----CCCACTTCTGGCGGTGTCTCTCCCTCCGACAGGGCCCTCAAGTTAC 1237

```

RESULT 11

```

US-09-925-065A-55409
/ Sequence 55409, Application US/0925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Mang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 55409
/ LENGTH: 1294
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-55409

```

Alignment Scores:

```

Pred. No.: 1,53e-13 Length: 1294
Score: 367.00 Matches: 143
Percent Similarity: 32.9% Conservative: 22
Best Local Similarity: 28.5% Mismatches: 86

```

```

Query Match: 12.3% Indels: 252
DB: 6 Gaps: 16
US-10-071-838-2 (1-549) x US-09-925-065A-55409 (1-1294)
Qy      202 ProGlnGlnAspAlaPheTrpAlaLeuValGlnLeuAlaSerGlnIleArgHisSerLeu 221
Db      34 CAGACCTCATAGGGAGAGACTGGCCAGAGCCGCACTTGAAGGGCTCAG---GAAAGCTTC 90
Qy      222 GlnGlyPheHisSerProAspGlnGlyThrValGlnIleGlnIleAspGlnIleHis 241
Db      91 AAGCCCTGGCAAGCCCTCT-----CTCCAGAGCCACATCCCCAC 132
Qy      242 -----ValValAlaThrSerGlnPro----- 248
Db      133 TCAATAGTCTCTCCATAGAGACTTCAAGACCTTGTCTGACCCGCTCTGGAGGG 192
Qy      249 -----LysThrMetGlyHisGlnAspLysAspLeuCyb 260
Db      193 CTCAGGGACCTCATAGGGGAGAGAGAGTCACTGACTCAGAGAACTG-AAAGCCCATGT 251
Qy      261 GlyGlnCybSerProLeuGly-----CysLeuIleArgIleLeuIle 274
Db      252 GCCCAGCTCGAGCCACAGCCCGCCTGGAGAGACCAAGGTCTTTCACACCTGTCTCC 311
Qy      275 AspGlyIleSerLeuGlnIleThrLeuArgLeuTrpAspValTyThrLeuValGlnGlyLeu 294
Db      312 CCAAGATCTCTTGGGCTCACCTCGCGCTGTGGACGTGTATCTGTATGTAAGCGAA 371
Qy      295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309
Db      372 CAGGCGTTATGCGCATACAGAAATCGCTTTAAGTTCAAGTAAAGTAACTACGTGT 431
Qy      309 ----- 309
Db      432 GCCCAGGGGCTGGGAGACCTGGGGGTGACACCCGACTGGCCGAGGGCAGCTTCT 491
Qy      309 ----- 309
Db      492 CACATGTCTCATGATCTCTGTCTGTGGCCAGAGGAGTGTGGCAGGTGGCTGGG 551
Qy      309 ----- 309
Db      552 CAGGACATGTGACACCGAGCCCATTTCCCACTAGACCAAGTAAAGTGAAGTGTGG 611
Qy      309 ----- 309
Db      612 TGAGCACTTCCCTGTCCAGATRCGCCCGACAGCAGTCTCTGTGTATGTGACGCC 671
Qy      309 ----- 309
Db      672 TGGGTGGCCACAAAAGATCCGGACCGCCAGTGGAGACTGAATGGCCACGGGGTA 731
Qy      310 -----LysAr 311
Db      732 TGAAGTGAACCATTTCCAGTAATCCCTGTGCTGATATCCACCTGTCCCTAAGCG 791
Qy      311 gleuthrlysrthrserargcybgllyprotrpalaargphecyasasnargphevalapth 331
Db      792 CCTCAGAGAGAGCTCAGAGTGTGGCCGCGGAGCACTTTTGGAACCGGTTGCTTATGC 851
Qy      331 rtrpalaargaspGlnAspThrValLeuLysHisLeuAlaGlnAlaSerMetLysLysLeuTh 351
Db      852 CTGGGCGAGGATGATGACACTGTGTCAAGACTTCTAAGGCTCTATAGAAAGAACTAAC 911
Qy      351 rArgLysGlnIlyAspLeuProProPro-AlaLysProGlnGlnIlySerSerAlaSerA 371
Db      912 AAGAAACAGGGGAGCTGCACCCAGGATG-----GGCTCCAGTGCATGT 959
Qy      371 rgrProValProAlaSerArgGlyIlyThrLeuCyblyGlyAspArgGlnAlaPro- 390
Db      960 CCCCT---CCCATGTCACTCTGGGGTACTCAATAGTACGGGA-----GTGC 1004

```



```

Db      1127 TACTGACATGCTGTGACGC----- 1147
Qy      471 HieTyRspPheArgInSerCySTrpValArgAlaIleSerGlnGluAspGlnLeuAla 490
Db      1148 -----AGGAGGGGAGATGTGGCGGAGA----- 1168
Qy      491 ProCyETrpGlnAlaGlnHisProAlaGluValArgSerAlaPheAlaProSer 510
Db      1169 ----- 1174
Qy      511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGluProCyValaProThr 530
Db      1175 -----AACAGGCCCT----- 1186
Qy      531 SerGlyProCyValeuCyGlyLeuHisLeuGlnSerGlnPheProGlyPhe 549
Db      1187 -----CCCACTTCTGCGGTGTCTCCCTCTCCCTCGCAGGGCCCTCCAAATTAC 1237

```

RESULT 13

```

US-10-301-480-156647
; Sequence 156647, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 10827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156647
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-156647

```

```

Alignment Scores:
Pred. No.: 1,53e-13      length: 1294
Score: 367.00      Matches: 143
Percent Similarity: 32.9%      Conservative: 22
Best Local Similarity: 28.5%      Mismatches: 86
Query Match: 12.3%      Indels: 252
DB: 9      Gaps: 16

```

US-10-071-838-2 (1-549) x US-10-301-480-156647 (1-1294)

```

Qy      202 ProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnArgHisSerLeu 221
Db      34 CCAGCTCATGGGACAGACTGGGCGCAGACCCGACTTGAGAGGGCTCAG--GGAAGCTTC 90
Qy      222 GlnGlyPheHisSerProArgGlnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnHis 241
Db      91 AAGCCCTGGGCAAGCCCTCT-----CTCCAGAGGACCAATCCCCAC 132
Qy      242 -----ValValAlaPheSerGlnPro----- 248
Db      133 TCMAATGAGTGCCTCCCATGAGAGCTTCAAGACCTTGTCTGACCCAGCGCTCTGAGAGG 192
Qy      249 -----LyThrMetGlyHisGlnAspIlyAspLeuCyS 260
Db      193 CTCAGGCGACCTCATGGGAGAGAGTCACTGACTCAGAGACTG--AAGCCCACTGT 251
Qy      261 GlnGlnCySerProLeuGly-----CyleuIleArgIleLeuIle 274
Db      252 GCGCACTCGAGCGACCAAGCCCGCTGAGAGACCAAGTCTTTCACACCTGCTGCC 311
Qy      275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlu 294

```

```

Db      312 CCACAGATCTCTCTCGGGCTCACCTCGCGCTGTGGAGCTGTATCTGTAGAAAGCGAA 371
Qy      295 GlnAlaLeuMetProIleThrArgIleAlaPheLeuValGln-Gln----- 309
Db      372 CAGGCGTGTATGCCGATTAACAAAGATCGCTTTAAGTTCAAGTAAAGTAACTACGTGT 431
Qy      309 ----- 309
Db      432 GCCCAGCGGGCCTGGGGAGCCCTGGGGTTCAGACCCGACTGGCCGAGGGCAGCTTCT 491
Qy      309 ----- 309
Db      492 CACACTGTCTCATGATCTCTGTCTGTGGCCCAAGAGAGGTCTGGCCAGGTGGCTGGG 551
Qy      309 ----- 309
Db      552 CAGAGACTGTGTACACCGAGCCATTCCCCAGATGACCCAGATGAAAGTCGAGAGTGTGG 611
Qy      309 ----- 309
Db      612 TGAGCACTTCCCTGTCCAGATRCGCCCGACAGCTCTCTGTATATCTGAGCGCC 671
Qy      309 ----- 309
Db      672 TGGGTTGGCCCAAAAGATCCGGCACCCGCCAGTGGAGACTGAAGTGGCCACGGGCTA 731
Qy      310 -----LysAr 311
Db      732 TGAGCTGTGACCATTCACAGGTATCCCTCGGCTGTATATCCACCTGTCCCTAAGCGG 791
Qy      311 GlnLeuThrIleThrSerArgCyGlyProTrpAlaArgPheCyAspArgPheValAspThr 331
Db      792 CTCACAAAGACTCTCAGAGTGTGGCCGTGGGACCTTTTGGAAACCGGCTTCTATATGC 851
Qy      331 rTPAlaArgAspGluAspThrValLeuIlyHisLeuArgAlaSerMetIlyAspLeuThr 351
Db      852 CTGGGCGCAGGATGATGACACTGTCTCAAGCACTTGAAGGCTCTATAGAAACTTAC 911
Qy      351 rArgIlyGlnGlyAspLeuProProPro-AlaIlyBProGlnGlnIlySerSerAlaSerA 371
Db      912 AAGAAAGCAGGGGAGACTGCCAACCCCAAGTG-----GGCTCAAGTGCATGT 959
Qy      371 rGProValProAlaSerArgIlyGlyIlyThrLeuCyIlyGlyAspArgGlnAlaPro 390
Db      960 CCCCT---CCCATGTCAACCTCTGGGTAGTCAAATGTAGGGA-----GTGC 1004
Qy      391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPro 410
Db      1005 CCGGAGCCGCAACCTTACTCTGGGCT-----TCTCTTCACTTTTCTTCTCT 1055
Qy      411 ArgSerSerThrProCyAspProGlyIlyAlaValArgGluAspThrTyrProValGlyThr 430
Db      1056 -----CCTTCTCTCTGTGACr-CTAAGAAATACAGAGGCCACCGGTCTCT 1102
Qy      431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450
Db      1103 CAGGCGAGCGCTCAAGTGGCT-----GTA 1126
Qy      451 GlnTrpAspSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470
Db      1127 TACTGACATGCTGTGACGC----- 1147
Qy      471 HisTyRspPheArgInSerCySTrpValArgAlaIleSerGlnGluAspGlnLeuAla 490
Db      1148 -----AGGAGGGGAGATGTGGCGGAGA----- 1168
Qy      491 ProCyETrpGlnAlaGlnHisProAlaGluValArgSerAlaPheAlaProSer 510
Db      1169 -----CCCTCC 1174
Qy      511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGluProCyValaProThr 530

```

Db 1175 -----AACAGCCCCCT----- 1186

Qy 531 SerGlyProCysLeuCyGlyIleuHleuGluSerGlnPheProGlyPhe 549

Db 1187 -----CCCACTTTCTGGGGTGTCTCCCTCTCCCCCTCGAGGGCCCTCCAAATTAC 1237

RESULT 14

US-10-301-480-770055

Sequence 770055, Application US/10301480

Publication No. US20060057564A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

TITLE OR INVENTION: in the Human Genome

FILE REFERENCE: 108827.137

CURRENT APPLICATION NUMBER: US/10/301,480

PRIOR FILING DATE: 2002-11-21

PRIOR APPLICATION NUMBER: US 10/215,598

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: US 60/311,695

PRIOR FILING DATE: 2001-08-10

NUMBER OF SEQ ID NOS: 1226818

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 770055

LENGTH: 1294

TYPE: DNA

ORGANISM: Homo sapien

US-10-301-480-770055

Alignment Scores:

Pred. No.: 1,536-13 Length: 1294

Score: 367.00 Matches: 143

Percent Similarity: 32.9% Conservative: 22

Best Local Similarity: 28.5% Mismatches: 86

Query Match: 12.3% Indels: 252

DB: Gaps: 16

US-10-071-838-2 (1-549) x US-10-301-480-770055 (1-1294)

Qy 202 ProGluGluAerAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeu 221

Db 34 CCAGCCTCATGGGCGAGCTGGGCGCCAGAGCCGACTTGAGAGGGCTCAG--GGAAGCCTC 90

Qy 222 GlnGlyPheHisSerProGlnGlyIleThrValGlnGlyLeuGlnAspGlnGlnGlnHis 241

Db 91 AAGCCTGGGCAAGCCCTCT-----CTCCAGAGCCACATCCCAAC 132

Qy 242 -----ValValAlaThrSerGlnPro----- 248

Db 133 TCMAATGAGTGCTCCCATGAGAGCTTCAAGACCTTGTCTGAGACCCAGCCTCTGAGAGG 192

Qy 249 -----LysThrMetGlyHisGlnAspLysLysPheLeuCys 260

Db 193 CTCAGAGCAACCTCATGGGGAAGAGGTCACTGACTGAGAGACTG--AAGCCCAAGTGT 251

Qy 261 GlnGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274

Db 252 GCGCAGCTCGAGCCACAGCCCCCAGCCTGGAAGACCAAGTTCTTTCACACTGCTGTCC 311

Qy 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlu 294

Db 312 CCACAGATCTCTCTGGGGCTACCTGGGCGCTGGAGCGTGTATCTGTAGAAAGCGAA 371

Qy 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGln----- 309

Db 372 CAGGGGTGATGCGCATTAACAAGATCGCTTTAAGGTTCAGTAGATTAAGTCTACGTGT 431

Qy 309 ----- 309

Db 432 GCCCAGCGGGCGCTGGGAGACCTTGGGGGTCAAGACCCGACCTGGCCGAGGAGACTTCT 491

Qy 309 ----- 309

Db 492 CACACTGTCTCATGTATCTCTGTCTGGCCAGAGGAGGTCTGGCCAGGTGGGCTGGG 551

Qy 309 ----- 309

Db 552 CAGAGCACTGTGACACCGAGCCATTCGCCACATGACCAAGATGAAGAATCGAGAGTGG 611

Qy 309 ----- 309

Db 612 TAGACATTCCTCTGTCCAGATCACCCCCAGACAGACTCTGTGTATATCTGACGCC 671

Qy 309 ----- 309

Db 672 TGGGGTGGCCACAAAGATCCGGCACCCGCCAGTGGAGACTGAAGTGGCCACGGGGTA 731

Qy 310 -----LysAr 311

Db 732 TAGACTGTGACCATTCACAGTAATCTCCCTGGCTGATATCCACCTGTCCCTAGAGCG 791

Qy 311 GLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspTh 331

Db 792 CTTACAGAAACGTCACAGGTGTGGCCGCTGGGACGTTTGGAAACCGGTTCTGTGATGC 851

Qy 331 TTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuTh 351

Db 852 CTGGCCAGGAGATGATGACACTGTCTCAAGCATCTTAAGGCGCTTATGAAGAAACTAAC 911

Qy 351 rArgLysGlnGlyAspLeuProProPro-AlaLysProGluGlnLysSerSerAlaSerA 371

Db 912 AAGAAAGAGGGGAGCTGCCACCCCAAGGTG-----GGCTCCAGTGCCTATGT 959

Qy 371 rArgProValProAlaSerArgGlyLysThrLeuCysLysGlyAspArgGlnAlaPro- 390

Db 960 CCCT-----CCCATGTACCTCTGGGGTATGCAATGATAGAGGA-----GTGC 1004

Qy 391 ProGlyProProAlaArgPheProArgProIleTrpSerLysSerProProArgAlaPro 410

Db 1005 CCGGAGACCCCAACCTACTACTCTGGGCT-----TCTCTTCACTCTTCTTCTCT 1055

Qy 411 ArgSerSerThrProCysProGlyLysValAlaValArgGluAspThrTyrProValGlyThr 430

Db 1056 -----CTCTTCTCTCTGAGACT-CTAAGAAAGTACAGAGAGCCACCGGTCT 1102

Qy 431 GlnGlyValProSerProAlaLeuAlaGlnGlyLysProGlnLysSerTrpArgPheLeu 450

Db 1103 CAGGCGACGGGCGCTCAGGCGT-----GTA 1126

Qy 451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470

Db 1127 TACTGACATGCTGTGCAGCC----- 1147

Qy 471 HisTrpAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGlnAspGlnLeuAla 490

Db 1148 -----AspAGGGGAGATGTGGCCAGA----- 1168

Qy 491 ProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaAlaProSer 510

Db 1169 -----CCCTCC 1174

Qy 511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530

Db 1175 -----AACAGCCCCCT----- 1186

Qy 531 SerGlyProCysLeuCyGlyIleuHleuGluSerGlnPheProGlyPhe 549

Db 1187 -----CCCACTTTCTGGGGTGTCTCCCTCTCCCCCTCGAGGGCCCTCCAAATTAC 1237

RESULT 15

US-10-301-480-770056

Sequence 770056, Application US/10301480

Publication No. US20060057564A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; TITLE OF INVENTION: in the Human Genome
 ; FILE REFERENCE: 108827.137
 ; CURRENT APPLICATION NUMBER: US/10/301,480
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/215,598
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 1226818
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 770056
 ; LENGTH: 1294
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-10-301-480-770056

Alignment Scores:
 Pred. No.: 1.53e-13 Length: 1294
 Score: 367.00 Matches: 143
 Percent Similarity: 32.9% Conservative: 22
 Best Local Similarity: 28.5% Mismatches: 86
 Query Match: 12.3% Indels: 252
 Gaps: 16

US-10-071-838-2 (1-549) x US-10-301-480-770056 (1-1294)

QY 202 ProGluGluAspAlaPheTTPAlaLeuValGlnLeuAlaSerGluArgHisSerLeu 221
 Db 34 CCGAGCTCATGGGCGAGCTGGGCGAGGACCCGACTTATAGAGGCTCAG--GGAAAGCTC 90
 QY 222 GlnGlyPheHisSerProAlaGlnGlyThrValGlnGlyLeuGlnAspGlnGlnGlnHis 241
 Db 91 AAGCCCTGGGCAAGCCCT 132
 QY 242 -----ValValAlaPheSerGlnPro----- 248
 Db 133 TCAATAGTGTCTCTCCATGAGAGGCTTCAAGACCTTGTCTGACCCAGCGCTCGAGGG 192
 QY 249 -----LysThrMetGlyHisGlnAspLysValAspLeuCys 260
 Db 193 CTCAGGCGACCTCATGGGAGAGAGAGTCACTGACTCAGAGAGCTG--AAGCCCACTGT 251
 QY 261 GlnGlnCysSerProLeuGly-----CysLeuLeuArgGlnLeuIle 274
 Db 252 GCGCACTCGAGGCAACGAGCCCAAGCTGAGAGAGCAGGTTCTTTTACACCTGCTGCC 311
 QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTyrAspValTyrLeuValGlnGlyGln 294
 Db 312 CCACAGATCTCTCTCGGGCTCACCTGGCGCTGAGGAGCTGTATCTGTATGAGAGGGA 371
 QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309
 Db 372 CAGGCGTGTATGCGATACAGAAATCGCTTAAAGTTCAGTAGAGTAAGTACGTGT 431
 QY 309 ----- 309
 Db 432 GCCCAGCGGGGCTGGGAGACCTTGGGGTCAAGCCCGAGTGGCCGAGGCGACTTCTCT 491
 QY 309 ----- 309
 Db 492 CACACTGTCTCATGATCTCTGTCTGTGGCCAGAGGAGGTCTGGCAGGTGGGCTGGG 551
 QY 309 ----- 309
 Db 552 CAGGACACTGTGACACGAGCCCATTCGCCCATGACCCAGATGAAAGTCGAGAGTGTGG 611
 QY 309 ----- 309
 Db 612 TAGACACTTCCTGTGCAATGATCCGCCCAAGCAGCTCTCTGTGTATATCTGAGCGCC 671
 QY 309 ----- 309
 Db 672 TGGGCTGGCCAGAAAGATTCGGGACCGCCAGTGGAGAGACTGAAGTGGCCAGGGGGTA 731

QY 310 -----LysArg 311
 Db 732 TGAGCTGTGACCATTCACAGTAATCTCCCTGGCTGTATATCCACCTGTCCCTAGAGCG 791
 QY 311 GlnLeuThrLysSerArgCysGlyProTTPAlaArgPheCysAsnArgPheValAspThr 331
 Db 792 CTTCAAGAAACCTCCAGGTGTGGCCCTGGGACGTTTGGAAACCGGTTCTGTATGCC 851
 QY 331 TTPAlaArgAspGlnAspThrValLeuLysHisLysValArgAlaSerMetLysLysLeuThr 351
 Db 852 CTGGGCCAGGATGATGACACTGTGTCTCAAGCATCTTAAGGCGCTCTATGAAGAAACTTAC 911
 QY 351 TArgLysGlnGlyAspLeuProProPro-AlaLysProGlnGlnGlySerSerAlaSer 371
 Db 912 AAGAAAGCAGGGGACCTGACCCCAAGGTG-----GGCTCAAGTGCCTTGT 959
 QY 371 rGProValProAlaSerArgGlyLysThrLeuCysLysGlyAspArgGlnAlaPro- 390
 Db 960 CCCCT--CCCATGTCACTCTGGGGTATGATCAATAGTAGGGA-----GTGC 1004
 QY 391 ProGlyProProAlaArgPheProArgProIleTyrSerAlaArgProArgAlaPro 410
 Db 1005 CCGGAGACCGGCAACCTACTACTGGGCT-----TCTCTTCACTTTTCTTCTCT 1055
 QY 411 ArgSerSerThrProCysProGlyValAlaValArgGlnAspThrTyrProValGlyThr 430
 Db 1056 -----CCTTCTCTCTGAGCTCTAAGAAAGTACAGAGGACCCAGCGTCT 1102
 QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTyrAspPheLeu 450
 Db 1103 CAGGGCAGCGCTCAGTGCCT-----GTA 1126
 QY 451 GlnTyrAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTyrPheArg 470
 Db 1127 TACTGGACATGCTGTGCACGC----- 1147
 QY 471 HisTyrAspPheArgGlnSerCysTyrValArgAlaIleSerGlnGlnAspGlnLeuAla 490
 Db 1148 -----AGGAGGGGAGATGTGGGAGAG----- 1168
 QY 491 ProCysTyrGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAlaIleProSer 510
 Db 1169 -----CCCTCC 1174
 QY 511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530
 Db 1175 -----AACAAGCCCT----- 1186
 QY 531 SerGlyProCysLeuCysGlyLeuHisLysGlnLysSerSerGlnPheProGlyPhe 549
 Db 1187 -----CCCACTTTCGCGGTGTCTCTCCCTCCGAGGGCCCTCCAGATGATC 1237

Search completed: April 5, 2006, 15:11:30
 Job time : 554 secs

This Page Blank (uspto)

Thu Apr 6 16:21:15 2006

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:54:53 ; Search time 1234 Seconds

(without alignments)
3679.002 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWAQAEREDIM.....TSGPCICGHLHSSQPPPGF 549

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_model -DEV=xlh
-Q=/abs/ABSSWB.spool/US10071838/runat_05042006_141920_5288/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=trnpsm -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abseq7
-USER=US10071838 @CGN 1.1 1026 @runat_05042006_141920_5288 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	5	US-10-071-838-1
2	2978	99.8	1993	8	US-10-835-882-31
3	2907.5	97.5	2146	5	US-10-071-838-3
4	2850	95.5	2084	7	US-10-302-172-692
5	2827	94.8	2072	7	US-10-302-172-691
6	2827	94.8	2647	6	US-10-108-260A-676
7	2784	93.3	1752	6	US-10-094-466-61

8	2773	93.0	1862	5	US-10-071-838-5
9	2191.5	73.5	7856	9	US-10-450-763-21335
10	2191.5	73.5	7856	9	US-10-450-763-24017
11	2191.5	73.5	7878	3	US-09-962-436-562
12	2191.5	73.5	7878	9	US-10-843-641A-3021
13	2191.5	73.5	7878	9	US-10-956-157-1259
14	2133.5	71.5	8180	9	US-10-450-763-24016
15	2133.5	71.5	8201	7	US-10-283-978A-272
16	2133.5	71.5	8201	8	US-10-278-698-640
17	2133.5	71.5	8201	8	US-10-278-698-1037
18	2133.5	71.5	8201	8	US-10-278-698-1037
19	2133.5	71.5	8201	8	US-10-278-698-1037
20	2133.5	71.5	8201	8	US-10-278-698-1037
21	1795.5	60.2	3304	9	US-10-450-763-21338
22	1795.5	60.2	3304	9	US-10-450-763-24018
23	1732	58.1	2838	9	US-10-450-763-1813
24	1732	58.1	2838	9	US-10-450-763-22833
25	1732	58.1	2838	9	US-10-450-763-22833
26	1533.5	51.4	2424	9	US-10-450-763-23327
27	1507.5	50.5	9805	3	US-09-764-869-1824
28	1507.5	50.5	9805	3	US-09-764-869-1824
29	1507.5	50.5	9805	3	US-09-764-869-1824
30	1507.5	50.5	9805	3	US-09-764-869-1824
31	1500.5	50.3	9404	3	US-09-764-869-1826
32	1500.5	50.3	9404	3	US-09-764-869-1826
33	1500.5	50.3	9404	3	US-09-764-869-1826
34	1500.5	50.3	9404	3	US-09-764-869-1826
35	1469	49.2	9804	3	US-09-764-869-1829
36	1469	49.2	9804	3	US-09-764-869-1829
37	1469	49.2	9804	3	US-09-764-869-1829
38	1469	49.2	9804	3	US-09-764-869-1829
39	1465	49.1	9796	3	US-09-764-869-1825
40	1465	49.1	9796	3	US-09-764-869-1825
41	1465	49.1	9796	3	US-09-764-869-1825
42	1465	49.1	9796	3	US-09-764-869-1825
43	1428.5	47.9	6696	3	US-09-764-869-1821
44	1428.5	47.9	6696	3	US-09-764-869-1821
45	1428.5	47.9	6696	3	US-09-764-869-1821

ALIGNMENTS

RESULT 1
US-10-071-838-2
Sequence, Application US/10071838
Publication No. US2003004481A1
GENERAL INFORMATION:
APPLICANT: Li, Jing
APPLICANT: Powers, Scott
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
FILE REFERENCE: 018781-007610US
CURRENT APPLICATION NUMBER: US/10/071,838
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: US 60/267,615
PRIOR FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1964
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human PRC17
NAME/KEY: CDS
LOCATION: (1)..(1650)
US-10-071-838-1
Alignment Scores: 2.19e-265 Length: 1964

Score: 2983.00 Matches: 549
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-071-838-2 (1-549) x US-10-071-838-1 (1-1964)

QY 1 MetcApValValGluValAlaGlySerTrpTrpAlaGlnGluValArgGluAspIleIleMet 20
DB 1 ATGGACCTGGTAGAGGCTCGCGGCACTTGGTGGCAACAAGCCAGAGCATATATG 60
QY 21 LysTrGluLysGlyGlyIleArgAlaGlyLeuProGluAspGlyProLysProPheArg 40
DB 61 AAATACGAAAGGAGACCGGAGCTGGCTGCAGAGGACAAAGGGGCTTAAGCTTTTGA 120
QY 41 SerTrpAsnAsnValAspPheLeuGlyIleValHisGluTrpGluLeuProLeu 60
DB 121 AGCTACAAACAACACGTCGATCTTGGGATTTGTACATGAGCGAGCTGCCCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 181 ACTGCGGGAGGCGGAGCAAAATTCGCGGAGATCCGCAAGAGCAAGTGGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTrpLys 100
DB 241 ATGCTGGAGACTCGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCTGACAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 301 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCCAGAACATTTGAGGAATG 360
QY 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 361 AAGTTGAAAACCCCGAAGATACCAATCATCAAGAGAGAGGCAAGGTCACTCGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 421 CACATCCAGGCAATCGACCGGAGCGTAAGCGGACATTAAGAGAGCATATATTTCCAGG 480
QY 161 AspArgTrpGlyTrpTrpLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
DB 481 GATCGAATACGAAACCAAGCAGCGGGAATCTCCACATCTCTCCGATATGAGAGATAT 540
QY 181 AsnProGluValGlyTrpArgAspLysSerHisIleAlaIleAspLeuLeuTrp 200
DB 541 AACCCGAGGTGGCTCTGCGAGGACCTGAGCCATCGCGCTTGTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 601 CTTCCTGAGGAGATGATCTCTGGGCACTGGTGCAGCTGGCTGGCAATGAGAGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 661 CTGAGAGGATTTCAACGCCAAAGGCGGAGCTCTCCAGGGGCTCCAAAGCAACAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 721 CATGTGGTAGCAAGTCAACCCCAAGACATGGGGATCGAGCAAGAAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 781 GGGCAGGTTCCTCCGTTAGGCTGCTCATCCGATATATTGATGACGGGATCTCTCTGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTrpLeuValGluGlyGluGlnAlaLeuMetProIle 300
DB 841 CTCACCTCTGGCGCTGGTGGAGCTGTATCTGTATACAAAGCGCAAGCGGTGATGCCATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTrpSerArgCysGlyPro 320
DB 901 ACAGAAATCCGCTTTAAGTTCAAGCAAGCGCTTCCAGAAAGCTCCAGGTGGGCGG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340

DB 961 TGGGCACTGTTTGGCAACCGGTTCTGTGTATACCTGGGCCAGGGATAGGACATGTGCTC 1020
QY 341 LysHisLeuAspArgIleSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 1021 AACATCTTTAGGGCTCTTAAAGAACTTAACAAAGAAAGCGGGAGCTTGCCACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
DB 1081 GCCAAACCGAGACGAAGGTGTGTGGCATCCAGGCTGTGGCTTCAAGTGGCGGAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1141 ACCCTCTGCAGAGGGGACAGGCGGCGCTCCAGGCCCAACACCGGTTCCGCGGCC 1200
QY 401 IleTrpSerLysSerProProAlaArgAlaProArgSerSerThrProCysProGlyAla 420
DB 1201 ATTTGGTCAGCTTCCCGCCAGCGGCACTCTGTTCTTCCACACCTGTCTGTGGGGCT 1260
QY 421 ValArgGluAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCCGGGAAGACACTTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1321 GGAGACCTCAGGGGTTCTGGAGATTCCTGCACTGGAATCCATGCCCCGCTCCACAG 1380
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTrpAspPheArgGlnSerCysTrpVal 480
DB 1381 GACCTGACCTTAAGAGGCGCTTGTGTCTCGCATTAATGATTTCAAGACAGAGCTCTGGGTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1441 CGTGCAATATCCAGAGAGGACAGCTGGCCCTCTGTGGCAGGCTGAACCTGGCGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1501 CGGCTGAGATCGGCTTCTGCTGCACAGACATGATTCGACACGGGCACTCCCTTCAG 1560
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1561 GCTAGAGCAACAAGCGGTGTGCTCCACCTCAGGCGCTTGTGCTGTGCGCTCCACTTG 1620
QY 541 GlusSerGlnPheProProGlyPhe 549
DB 1621 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1647

RESULT 2
US-10-839-882-31
; Sequence 31, Application US/10839882
; Publication No. US20040203106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: CORLEY, Neil C.
; APPLICANT: LAL, Preeti
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: JUNMING, Yang
; APPLICANT: SHIH, Leo L.
; TITLE OR INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
; FILE REFERENCE: PF-0619 PCT
; CURRENT APPLICATION NUMBER: US/10/839,882
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US/09/807,452
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/175,737; unassigned; 60/118,559; 09/249,740; unassigned
; 60/154,336
; PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11;
; 1999-04-22

NUMBER OF SEQ ID NOS: 44
 SOFTWARE: PERL Program
 SEQ ID NO: 31
 LENGTH: 1993
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No: 4325626CBI
 US-10-839-882-31

Alignment Scores:

Pred. No.:	6.45e-265	Length:	1993
Score:	2978.00	Matches:	548
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	1
Query Match:	99.8%	Indels:	0
DB:	8	Gaps:	0

US-10-071-838-2 (1-549) x US-10-839-882-31 (1-1993)

QY 1 MetAapValValGluValAagLysSerTPTPPALagInGluArgGluAapIleIleMet 20
 DB 42 ATGCACTGTGTAGAGGTCGGCGGCAAGTGTGTGGGCAAGAGCGAGAGGACATCATTTATG 101
 QY 21 LysTyrGluLysGluHisArgAlaGlyLeuProGluAapLysGlyProLysProPheArg 40
 DB 102 AAATACGAAAGAGACACCGAGCTGGCTGCAGAGAGACAGAGGCGCTTAAGCTTTTCA 161
 QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
 DB 152 AGCTACAAACACAGCTGCATTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 221
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluLysSerArgLysSerLysTyrValAap 80
 DB 222 ACTGGCGGAGCGGAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 281
 QY 81 MetLeuGlyAapTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
 DB 282 ATGCTGGAGACTGGAGAAATACAAAGACAGAAAGCTCATATGATGACCTGACAG 341
 QY 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
 DB 342 GGAATGCCCATGAACATCCGGGGCCGATGTGTCTGCTCTGAACTGAGGAAATG 401
 QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 DB 402 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGAGGAGGATCATCTAG 461
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 DB 462 CACATCCAGCGCATGACCGGAGCGGAACTACTCCACATCTCTGGCAATAGAGATAT 521
 QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
 DB 522 GATCATACGGAACCAAGAGCGGAACTACTCCACATCTCTGGCAATAGAGATAT 581
 QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleValAlaLeuPheLeuTyr 200
 DB 582 AACCCGAGGTGGGTACTGACGAGGACCTGACCATGCGCCCTGTGTTCTCTCTAT 641
 QY 201 LeuProGluGluAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 DB 642 CTTCCTGAGAGAGATGCTATCTGGGACCTGTGACGTGCTGGCCAGTGAAGGACCTCC 701
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
 DB 702 CTGCAGGAGATTCAACGCCCAATGCGGAGACCTGTCCAGGGGCTCCAAAGACCAAGAG 761
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 260
 DB 762 CATGTGTAGACCACTCAACCAAGCAATGAGGAGCATCAAGACAAAGAAATCATATGT 821

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
 DB 822 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATATTTGACGGGATCTCTCGGG 881
 QY 281 LeuThrLeuArgLeuTyrAspValTyrLeuValGluGlyGluAlaLeuMetProIle 300
 DB 882 CTCACCTTCGCTGTGGGACCTGTATCTGTAGAGAGCGAAGACGCTGTGATGCCGAT 941
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 DB 942 ACAAGATTCCTTTAAGGTTACAGAGAGAGCCCTCCAGAAACGTCAGGCTGCGCCG 1001
 QY 321 TTPAlaArgPheCysAsnAspPheValAspTyrTTPAlaArgAspGluAspHisValLeu 340
 DB 1002 TGGGACGTTTGTGAAACCGGTTCTGTATACCTGGGCGCAGAGATGAGACATGTGCTC 1061
 QY 341 LysHisIleLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
 DB 1062 AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1121
 QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
 DB 1122 GCCAAACCCGAGCAAGGCTGTGCGCATCCAGGCTGTGCGGCTTCACTGCGGGAAG 1181
 QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 DB 1182 ACCCTCTGAGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1241
 QY 401 IleTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyLysAla 420
 DB 1242 ATTTGGTACGTTCCCGGCAAGGAGCCTGTTCTTCCACACCTGCTGTGTGGGCT 1301
 QY 421 ValArgLysAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 DB 1302 GTCCCGAAGACACCTTACCTGTGGGCACTCAGGGGTGTGCCAGCCGCGCTGCTCAG 1361
 QY 441 GlyLysProGlnGlySerThrArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
 DB 1362 GAGGAGCTCAGAGGTTCTCGAGATTCCTGAGTGAATCTCATGCTCCGCTCCAGAG 1421
 QY 461 AspLeuAspValGluGlyProTTPPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
 DB 1422 GACCTGAGAGTGAAGAGGCTTGTGCTCCGCTATGATTTCAACAGAGCTGCTGGCTC 1481
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTTPGlnAlaGluHisProAlaGln 500
 DB 1482 CGTGCATATCCAGAGAGACAGCTGGGCTGCTGAGAGGCTGAGAACACCTTCGAG 1541
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 DB 1542 CGGTGAGATTCGCTTGTGCTGACACCAAGCACTGATTCGACAGAGGACACCTTCAGA 1601
 QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisIle 540
 DB 1602 GCTAGGAGAGAAACAGCCTGTGCTCCACCTCAAGGCTTGTGCTGCGGCTCCACTTG 1661
 QY 541 GlnSerSerGlnPheProProGlyPhe 549
 DB 1662 GAAAGTCTCAGTTCCTCCAGGCTTC 1688

RESULT 3
 US-10-071-838-3
 ; Sequence 3, Application US/10071838
 ; Publication No. US2003004481A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Jing
 ; APPLICANT: Powers, Scott
 ; APPLICANT: Xiang, Phil
 ; APPLICANT: Peng, Yue
 ; APPLICANT: Tularik Inc.
 ; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
 ; FILE REFERENCE: 018781-007610US
 ; CURRENT APPLICATION NUMBER: US/10/071,838

Qy 540 LeuGlusSerSerGlnPheProGlyPhe 549
Db 1800 TTGGAAAGTTCTCAGTTCCCTCCAGGCTTC 1829

RESULT 4
US-10-302-172-692
Sequence 692, Application US/10302172
Publication No. US20040053250A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and
FILE REFERENCE: 803 ICNCP
CURRENT APPLICATION NUMBER: US/10/302,172
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/225,251
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: PCT US02/05095
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 09/799,451
NUMBER OF SEQ ID NOS: 950
SOFTWARE: pc FL_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163)..(1755)
US-10-302-172-692

Alignment Scores:
Pred. No.: 4,58e-253 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.4% Conservative: 1
Best Local Similarity: 96.2% Mismatches: 2
Query Match: 95.5% Indels: 18
DB: 7 Gaps: 1

US-10-071-838-2 (1-549) X US-10-302-172-692 (1-2084)

Qy 1 MetLaPvAlaValGluValAlaGlySerTTPTPAlaGlnGluArgLysAlaMet 20
Db 163 ATGAGCGTGTGAGAGGTCCGGGCACTTGTGGGCAAGAGGAGCAATCATTAAG 222

Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATTCGAAAGGAGACCGGAGCTGGCTGCGAGAGCAAGGGGCTTAAGCTTTTCA 282

Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheArg 60
Db 283 AGCTACACCAACCAAGCTGATCATTTGGGATTTGATCATGACGAGCTGCTCTCG 342

Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerTyrValAsp 80
Db 343 ACTGGCGGAGGAGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 402

Qy 81 MetLeuGlyAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGGAGAAATACAAAGCAAGAGAAAGCTCATATGAGCGTCAAG 462

Qy 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGATCATCCGGGGCCGAGTGTGCTCTCTGAAACACTAGAGAAAG 522

Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528

Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160

Db 529 AACATCAGCGCATGACCGGAGCTTAAGCGGACATTAAAGAAAGATATTTCTTCAG 588
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 589 GATCGATACGAAACCAAGCAGCGGGAATCTTCAATCTCTCTGCAATATGAGAGTAT 648
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
Db 649 AACCCGAGGTGGCTACTGACAGGAGCATGAGCAATGCGCGCTTGTTCTCTCTAT 708
Qy 201 LeuProGluGluAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCTGAGAGAGATCATTTCTGGGACTGTGTGACCTGTGCGCATGAGAGCACTCC 768
Qy 221 LeuGlnGlyPheHisSerProAsnGlyIleThrValGlnGlyLeuGlnAspGlnGlu 240
Db 769 CTGACAGGATTTCAACGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCAAGAG 828
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCACAAACCCAGACCATGGGCAATCAGGAAAGAAATCTATGT 888
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGTGTTCCTCCGTTAGGCTGTCTCATCCGATTTGATGACGGAGTCTCTCGGG 948
Qy 281 LeuThrLeuArgLeuTTPAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 949 CTCACCTCGCGCTGTGGAGCGGTATCTGTATGAAGGCAAGCGGCTTGATGCCGATA 1008
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATGCTTTAAAGTTTACAGCAAGACCCCTCAAGAAAGCTGAGTGTGCCCCG 1068
Qy 321 TTPAlaArgPheCysAsnArgPheValAspTTPAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGCACTTTTTCACCGGTTCTGTATACCTGGGCCAGAGATGAGCACTGTGCTC 1128
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1129 AAGCATCTTAAGGCGCTTATGAAAGAACTTAACAAGAAAGCAGGGGAGCTGCAACCCCA 1188
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1189 GCCAAACCCGACCAAGGCTGTGCGCATCCAGGCTGTGCTTCACTGCGGAGAG 1248
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGluProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGAGCAGGAGGCTCTCAAGGCCACAGCCCGGTTCCGCGGCCC 1308
Qy 401 IleTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
Db 1309 ATTGTGTAGCTTCCCGGCAAGGCACTGCTTCTTCCACACCTGTCTGTGGGGCT 1368
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACACTTACCTGTGGGCACTCAGGGGTGTGCCAGGCCGCTGTGCTCAG 1428
Qy 441 GlyLysProGlnGlySerTTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
Db 1429 GAGGAGCTCAGGGTCTCGAGATTCCTGAGTGAATTCATGATCCATGCCCCGCTCCACAG 1488
Qy 461 AspLeuAspValGluGlyProTTPPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
Db 1489 GACCTGGAGCTGAAGAGGCTTGTTCGCACTTATGATTTCAAGACAGGCTGTGGCTC 1548
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTTPGlnAlaGluHisProAlaGlu 500
Db 1549 CGTGCATATCCACAGAGAGACAGCTGCGCTCTCTGCGAGGCTGAACACCTTCGAGAG 1608
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520

Db 1609 CGGGTGAATCGCTTTCGTGACCCAGACAGTATTCGACAGGGACCCCTTTCAGA 1668
Qy 521 AAlarngpuglunpocysValaProThrSerGlyProCysLeuCyseGlyLeuHisLeu 540
Db 1669 GCTAGGAGACGAAACCGGTGTCTCCACCTCAGAGGCTTGTCTGTGGCTCCTCAGCTTG 1728
Qy 541 GluSerSerGlnPheProGlyPhe 549
Db 1729 GAAAGTCTCAGTTCCTCCAGGCTTC 1755
RESULT 5
US-10-302-172-691
; Sequence 691, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radcoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803.1CNCB
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: PC_FL_genes Version 2.0
; SEQ ID NO 691
; LENGTH: 2072
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)..(1743)
US-10-302-172-691
Alignment Scores:
Pred. No.: 6,08e-251 Length: 2072
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 94.8% Mismatches: 2
Query Match: 7 Indels: 22
Gaps: 1
US-10-071-838-2 (1-549) x US-10-302-172-691 (1-2072)

Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetIleGlyLysGlyLysArgSerSerGly 140
Db 523 AAGTTGAAAACCCCGGAGATACAGATCATATAGAGAGAGGAGGACATCATCTAG 582
Qy 141 HisIleGlnArgIleAspArgValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 583 CACATCCAGGCGATCGACCCGGAGCTTAAAGGGAGCATTTAAGAAAGCATATATCTTCAGG 642
Qy 161 AspArgTyrGlyThrLysGlnArgGlyLeuLeuHisIleLeuLeuAlaTyrGlnGlyTyr 180
Db 643 GATCGATACGGAACCAAGACCGGGAACCTTCTCAGATCTCTCGGATATGAGAGAT 702
Qy 181 AsnProGlyValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
Db 703 AACCCGAGGTGGGCTTACTGACAGGACCTGAGCAATGCGCCGCTTGTCCTCTCTAT 762
Qy 201 LeuProGlyLysLysAspArgPheThrAlaLeuValGlnLeuLeuAlaSerGlyLysLeu 220
Db 763 CTTCCTGAGAGAGATGATCTGTGGCACTGTGCTGAGCTGTGCTGAGTGAAGGCACTCC 822
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 823 CTGAGGAGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCAACAGAG 882
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 883 CATGTGTGACCAAGTACACCAACCAAGACCATGGGGATCAG----- 924
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCTCGGG 936
Qy 281 LeuThrLeuArgLeuTyrPheValGlyLeuValGlyGlnAlaLeuMetProIle 300
Db 937 CTGACCTCGGCTGTGGAGCTGTATCTGTGAAAGGCAACAGGCTTTGATGCCGATA 996
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACAAGATCGCTTTAAGTTACAGAGAGCGCTCAGAAAGCTCAGAGTGTGGCCG 1056
Qy 321 TyrAlaArgPheCysAsnArgPheValAspThrTyrAlaArgAspGlyAspThrValLeu 340
Db 1057 TGGCGAGTTTGTGCAACCGGTTCTGTGATACCTGGCCCGGATGAGACACTGTGCTC 1116
Qy 341 LysHisLeuArgLysSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1117 AAGCATCTTAGGGCTTATGAAAGAACTTAAAGAAAGAGGGGAACTGCAACCCCA 1176
Qy 361 AlaLysProGlyGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1177 GCCAAACCGAGAGAGGTGTGGCATTCAGAGGCTGTGCGGCTTCACTGGGGAGAG 1236
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTTGCAAGGGGAGCAAGGCAAGCCCTTCCAGGCGCCCAAGCCCGGCTCCGCGGCC 1296
Qy 401 IleThrSerLysSerProProArgLysProArgSerSerThrProCysProGlyValAla 420
Db 1297 ATTGTGTACCTTCCCGGCAAGGCAAGCTGTCTTCCACACCTCTCTCTGTGGGGCT 1356
Qy 421 ValArgGlyLysAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1357 GTCCGGGAGACACCTTACCTGTGGGACATCAAGGTGTGCCAGCCGCGGCTGTGGCTCAG 1416
Qy 441 GlyLysProGlnGlySerTyrPheLeuGlnTyrAsnSerMetProArgLeuProThr 460
Db 1417 GAGAGACCTCAGGGTTCTGAGAGATCTCTCAGTGGAACTCATGCCCCCTCCCAAG 1476
Qy 461 AspLeuAspValGlyGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
Db 1477 GACCTGGACGTAGAGGGCCCTGTGTTCCGCAATTATGATTTCAAGACAGACTGTGGGCTC 1536

Db 2189 CGGGTGAGATCGGCTTTGCTGTGACCCAGCACTGATTCGACAGGAGCCCTTCAG 2248
Qy 521 AAlatgarpGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db 2249 GCTAGGGAGCAAGCCGCTGTGCTCCACCTCAGGGCCCTTCGCGGCTCCACTTG 2308
Qy 541 GluSerSerGlnPheProGlyPhe 549
Db 2309 GAAGTCTCACTTCCTCCAGGCTTC 2335
RESULT 7
US-10-094-466-61
; Sequence 61, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Strytek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 61
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)..(1678)
US-10-094-466-61
Alignment Scores:
Pred. No.: 4,67e-247 Length: 1752
Score: 2784.00 Matches: 515
Percent Similarity: 96.28 Conservative: 13
Best Local Similarity: 93.88 Mismatches: 21
Query Match: 93.34 Indels: 0
Gaps: 0
US-10-071-838-2 (1-549) x US-10-094-466-61 (1-1752)

Db 151 AGCTACAACAACGTCGATCTTTGGGATTGTACAGAGACGGGCTCCCTCG 210
Qy 61 ThrAlaArgGluAlaValGlnIleArgArgGluIleSerArgLysSerIleTPValAsp 80
Db 211 ACTGGCGGAGGTGAAGCAAAATTCGGCGGAGATCCGGAAGACCAAGTGGGTAA 270
Qy 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 271 ATGCTGGGAGAAATGGGACACCTTCGAAAAACGCAAGAAAGCTCATAGATCGAGGTAC 330
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 331 GGAATTCCTCAATGAACATCCGGGGCCCGATGTGTCAGTCCCTCGAACATTGGGAATC 390
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysAspSerSerGlu 140
Db 391 AAGTTGAAAAACCCGGAAGATACCGATCATAGAGAGAGGACAGAGGTCATCTGAA 450
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 451 CACATCCAGCAGATGACCTGAGCTGAGGTCGAGGACATTAGAGGACATATCTTCAGG 510
Qy 161 AspArgTyrGlyThrLysGlnArgLysLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 511 GATCATACCGAAACCAAGCAGCGGAACTATTACCTCTCGCGGTATGAGAGTAT 570
Qy 181 AsnProGluValGlyTyrCysArgAspLysSerHisIleAlaAlaLeuPheLeuTyr 200
Db 571 AACCCGAGGTGGCTACTGAGGACCTGAGCCACATCCGCGCTTGTCTCTTAT 630
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuLysSerGluLysSer 220
Db 631 CTTCCTAGAGAGATGATTCATTCGGGACCTGGTGCAGCTGTGGCAGTAGAGGACATCC 690
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 691 CTGCAGGGTTTTCACACCCCAAAATGGCGGACCGCTCCAGGGGCTCCCAAGCAACAGAG 750
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 751 CATGTGGTACCCACGTAACCTACCAACACATGGGCATCAGACAAAGAAATCTATGT 810
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 811 GGGCAGTGTGTCCTTAAGGTGCTGCATCGGATATTATGACGGGATCTCTCGGG 870
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 871 CTCACCTCGCGCCGTGGAGCTGTATCTGTAGAAAGCAAGCGGTGATGCGGATG 930
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 931 ACAAAGATCGCTTAAAGTTCACTAGAGGCTTCAGAAAGCTCAGAGTGTGGCCG 990
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 991 TGGGCAGTGTTTTGAACCGGTTGTGATGCTCGGSCAGGATGATGACACTGTCTC 1050
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1051 AAGCATCTTAGGGCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1110
Qy 361 AlaLysProGluGlnGlySerSerValSerArgProValProAlaSerArgGlyLys 380
Db 1111 GCCAAACCCGAGCAAGGATGTCGGCATCCAGGCTGTGCGAGCTTACGTCGGGAG 1170
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1171 ACCCTCTGCAAGGGGACAGGCAAGCCCTCAGGCCCAACAGCCGGTCCCATGGCCC 1230
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420

Db 1231 ATTGGTCAGCTTCCCGCCAGCGGCACTGCTCTTCCACACCTGCTGCGGGCT 1290
 Qy 421 ValArgGlnAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1291 GTCGGGAAAGACACCTTACCTGTGGCACTCAGGGGTGTGCGAGCCGCGCTGCTCAG 1350
 Qy 441 GtGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProAlaGlyProThr 460
 Db 1351 GGAAGCACTCAGGGTCTCTGAGATTCCTGCAATGCAATGCAATGCAATGCAATGCA 1410
 Qy 461 AspLeuAspValGlnGlyProTyrPheArgHisTyrAspPheArgGlnSerGlyTrpVal 480
 Db 1411 GACCTGAGCGTACGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
 Qy 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
 Db 1471 CGGCAATATCCAGAGAGACAGCGGCGGCACTGCTGAGGCTGAGGCTGAGGCTGAG 1530
 Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerArgGlnGlyThrProPheArg 520
 Db 1531 CGGCTAGATCCGCTTCACTGCGCCAGACATGATTCGACCAAGGCGACCTCTTCAAG 1590
 Qy 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
 Db 1591 GCTAGGAGAGACAGAGAGTGTGCTCCACCTCAGACCTTGGCTGCGGCTTCCACTTG 1650
 Qy 541 GluSerSerGlnPheProProGlyPhe 549
 Db 1651 GAAAGTTCTCAGTTCTCCCTCAGGCTTC 1677

RESULT 8

US-10-071-838-5
 : Sequence 5 Application US/10071838
 : Publication No. US20030044814A1

: GENERAL INFORMATION:
 : APPLICANT: Li, Jing
 : APPLICANT: Powers, Scott
 : APPLICANT: Xiang, Phil
 : APPLICANT: Peng, Yue
 : APPLICANT: Tularik Inc.
 : TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
 : FILE REFERENCE: 018781-007610US
 : CURRENT APPLICATION NUMBER: US/10/071, 838
 : PRIOR FILING DATE: 2002-05-07
 : PRIOR APPLICATION NUMBER: US 60/267,615
 : NUMBER OF SEQ ID NOS: 18
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 5
 : LENGTH: 1862
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)..(1538)
 US-10-071-838-5

Alignment Scores:
 Pred. No.: 5, 22e-246 Length: 1862
 Score: 2773.00 Matches: 515
 Percent Similarity: 93.8% Conservative: 0
 Best Local Similarity: 93.8% Mismatches: 0
 Query Match: 93.0% Indels: 34
 DB: 5 Gaps: 1

US-10-071-838-2 (1-549) x US-10-071-838-5 (1-1862)

Qy 1 MetAspValAlaGlnValAlaGlySerTrpPheArgGlnGlnGlnGlnGlnGlnGlnMet 20
 Db 1 ATGACGCTGTAGAGGTCCGCGGCAAGTGTGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 60

Qy 21 LysTyrGlnLysGlnHisArgAlaGlyLeuProGlnAspLysGlyProLysProPheArg 40
 Db 61 AAATACGAAAGGAGACAGAGCTGGGCTGCGAGAGACAGAGGAGGCTTACCTTTTCA 120
 Qy 41 SerTyrAsnAsnValAlaPheHisLeuGlyIleValHisGlyThrGlnLeuProProLeu 60
 Db 121 AGCTAACACAAACAAGCTCGATCTTTGGGGATTGTACATGAGACGAGAGCTGCTCTG 180
 Qy 61 ThrAlaArgGlnAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80
 Db 181 ACTGCGGAGGAGCGAAGAAATTCGCGGAGATCAGCGGAAAGAGCAAGTGGTGAT 240
 Qy 81 MetLeuGlnAspTrpGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
 Db 241 ATGCTGGAGAGCTGGAGAAATACAAAGACAGAGAACTCATTAATGAGGCTTCAAG 300
 Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGlnMet 120
 Db 301 GGAATCCCATGAAATCCGCGGCGCGATGTGGTCACTCTCTGAAATGAGAAATG 360
 Qy 121 LysLeuLysAsnProGlyValArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGln 140
 Db 361 AAGTTGAAAACCCCGAAGATACAGATCATGAAGAGAGAGCAAGAGTCACTGAG 420
 Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 421 CACATCCAGCGCATTCAGCGGAGCGTAAACCGGAGACCTTAAGAGAGATATTTCTTCAG 480
 Qy 161 AspArgTyrGlyThrLysGlnArgGlnLysLeuHisIleLeuLeuAlaTyrGlnGlnTyr 180
 Db 481 GATCGATACGGAACCAAGAGCGGGAATCATCTCCATCTCTCTGCAATGAGAGATAT 540
 Qy 181 AsnProGlnValAlaGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
 Db 541 AACCCGAGAGTGGGCTACTGAGGAGCGAGCATGACCATGCGGCTTGTCTTCTCTAT 600
 Qy 201 LeuProGlnGlnAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnArgHisSer 220
 Db 601 CTTCCTGAGAGAGATGATTCATTCGGCATCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGln 240
 Db 661 CTGCGAGGATTTCAAGCCCAATGCGGAGCGCTGCGAGGCTCCAAACCAAGAGAG 720
 Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 721 CATGTGTAGCCACGTCACCAACCAACCATGAGGAGCATCAG----- 762
 Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
 Db 762 ----- 762
 Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
 Db 763 -----TATCTGTGAGAAAGCGAACAAGCGCTTGTATGCCGATAT 798
 Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 799 ACAAGAAATCGCTTTAAGTTTACAGAGAGCGCTTCACAGAAAGCTCCAGGTGTGCGCG 858
 Qy 321 TrpAlaArgPheCysAsnArgPheValAspTrpTrpAlaArgAspGlnAspThrValLeu 340
 Db 859 TGGGCACTTTTGGAAACCGGTTCGTATATCTGGGCGAGGAGATGAGGACATGTGCTC 918
 Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
 Db 919 AAGCATCTTAAAGGCTCTATGAGAACTTAACAAGAAAGAGAGGAGGAGCTCCACCCCA 978
 Qy 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
 Db 979 GCCAAACCCAGAGAGGATCGCGGATTCAGAGGCTGTGCGGCTTCAAGTGGCGGAGAG 1038
 Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400

Db 1039 ACCCTCTGCAAGGGGACAGGCAAGCCCTCTCAAGCCACAGCCCGGCTCCCGGAGCC 1098
Qy 401 TLETPSERALASERPROPROAVALAPROARSESERThrProCybProGlyValA 420
Db 1099 ATTTGGTCAGCTTCCCGCCCAAGGCACTTCTTCCACACCTCTGCTGGGGCT 1158
Qy 421 ValArgLysArgPhePheProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1159 GTCCGGGAAGACACTACCTGAGGACCTAGAGGTGGCCAGCCGAGCCCTGCTCAG 1218
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1219 GAGAGACTCAAGGCTTCTGAGATTCTCTGAGTGAAGTCAAGCCCGGCTCCACAG 1278
Qy 461 AspleuAapValGluGlyProTrpPheArgHisTrpAsnPheArgGlnSerCybTrpVal 480
Db 1279 GACCTGGACGTAGAGGGCCCTTGGTCCGCAATTATGATTCAGACAGAGCTGCTGGGTC 1338
Qy 481 ArgAlaIleSerGlnGluArgGlnLeuAlaProCybTrpGlnAlaGlnHisProAlaGlu 500
Db 1339 CGTCCATATCCAGAGAGACACTGAGCCCTGCTGGCAGGCTGAACACCTTGGAG 1398
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1399 CGGGTAGATCGGCTTCTGCTGCAACCAAGCACTGATTCGACAGGGGACCCCTTACA 1458
Qy 521 AlaArgAspGluGlnProCybAlaProThiSerGlyProCybLeuCybGlyLeuHisLeu 540
Db 1459 GCTGGGACGMAACGCGGTGTCTCCACCTCAGGGGCTTGGCTGTGGGCTTCACCTTG 1518
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 1519 GAAAGTTCAGTTCCTCCACAGGCTTC 1545

RESULT 9

US-10-450-763-21335
Sequence 21335, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 21335
LENGTH: 7856
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (1697)..(4054)
OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number
OTHER INFORMATION: X63546, Smith-Waterman Score=423.
US-10-450-763-21335

Alignment Scores:

Pred. No.: 1.62e-191 Length: 7856
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
Gaps: 1

US-10-071-838-2 (1-549) x US-10-450-763-21335 (1-7856)

Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGlnLysArgGluAspIleIleMet 20
Db 1697 ATGACATGTTAGAGATGACATGATTTTCAGGACACAGAGCGGAAGACATACCTTATG 1756
Qy 21 LysTrpGluLysGlyHisValGalaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGAGACCCAGCTGGGCTGCCAGAGACAAAGGGGCTGAGCCGCTT --- 1813
Qy 41 SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGlnLeuProProLeu 60
Db 1814 GGAATCAACAGCAGCATGTATGCTTTGGCATTTTGCATGAGACGAGAGCTGCTCTGTG 1873
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGCAGGGAGGCGAAGAAATTCGGCGGGAATGACACAGACAGCAAGTGGATGGAA 1933
Qy 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTrpLys 100
Db 1934 ATCTGGAGAAATGGGAGACATATTAAGCACATGACAACTCATATGATCGATGTACAG 1993
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GGAATTCCTCAATGAACATCCGGGGCCCGGTGTGTCAGTCTCTTGAACATTCAGAAATC 2053
Qy 121 LysLeuLysAspProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAAAACCCCGGAAGATACAGATCATAGAGAGAGGGGCAAGAGTCACTCGAA 2113
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACACATCGACCTCGAGCGAGGACCACTCTCCGGAACCATGTCTTTTATGG 2173
Qy 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
Db 2174 GATCGATATGACCCAAAGCAGAGGAACTATTCTACATCTCTGAGCTTATTCGAGAT 2233
Qy 181 AspProGluValGlyTrpCybAspAspLeuSerHisIleAlaAlaLeuPheLeuLeuTrp 200
Db 2234 AACCCGAGGTGGGCTACTCTCAGAGACCTGAGCAATCAATCAGCCCTTGTTCCTCTTAT 2293
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCTTAGAGAGACGATCTGGGCACTGGTCACTGCTGGCCAGTGAAGGCACTCC 2353
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 2354 CTGCCAGGATTCACAGCCCAATGGGAGACGTCCAGGGGCTCCAAAGACAAACGAG 2413
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCyb 260
Db 2414 CATGTGTACCCAAAGTACAAACCAAGACCATGTGGCATGAGAACAAAGAGGTATATGC 2473
Qy 261 GlyGlnCybSerProLeuGlyCybLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTCTCGTTAGGCTGCTCTCCGAAACCTGATTAACGGGAGTCTCTTCGGG 2533
Qy 281 LeuThrLeuArgLeuTrpAspValTrpLeuValGluGlnAlaLeuMetProIle 300
Db 2534 CTACCTCTGGCCCTGTGGAGCGTATTTGTGGAGAAACAGGCTGTGATGCCAATA 2593
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCybGlyPro 320
Db 2594 ACCAGCATTTGCTCTTAAGTTTACAGCAGAACGCTCTCATGAAGCATCAAGGTGGCTG 2653
Qy 321 TrpAlaArgPheCybAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGCAGCTGTGGGAAACCAATCTTCGATACCTGGGCGCATGAACGATGACACCGTCTC 2713
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATTTAGGGCTCTTACGAAGAAACTGAACAAAGAAAGAGGGGACCTGCCACCCCA 2773

QY 361 AAlayProglunglInglYserSerAlaSerArgProValProAlaSerArglYglYlys 380
 Db 2774 GCCAAACGCGAGCAAGGGCTCTTGCCACCCAGGCTGTGCGGCTTACCGTGGGAG 2833
 QY 381 ThleuCyelyseglYasparGlnAlaProProglYProProAlaAargPheProArgPro 400
 Db 2834 ACCCTGCAAGGGGTATAGAGGGCCCTCCAGGCCACAGCCAGCTTCAGCGGCC 2893
 QY 401 lletPserAlaSerProProAlaProArgSerSerThProCysProglYglYAla 420
 Db 2894 ATTGCTGACCTTCCCGGCAATGGGCACTCTGTTTCCAGCCCTGTCTGTGGGGCT 2953
 QY 421 ValArgGluAspThrTyProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 2954 GTCCGGGAAGACAGTACCTGTGGGCACTCAGGGGTGTGCCAGCCCTGGGCTCAG 3013
 QY 441 GlyGlyProGlnInglYserTrpArgPheLeuGlnTrpAsnSerMetProAlaLeuProThr 460
 Db 3014 GGAGGACCTCAGGGGTCTCGAGATCTCTGGAGTGGAAAGCAATGCCCGCTCCCAAG 3073
 QY 461 AspleuAspValGluGlyProTrpPheArgHisTrpAspPheArgGlnSerCysTrpVal 480
 Db 3074 GACCTGATATAGGGGGCTTGTGATCTCCCATTTATGATTTTGAACGAGCTGTGGGTC 3133
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
 Db 3134 CGTGCAATATCCAGAGAGACCAAGCTGGCACCCTGCTGGAGGCTGAACCTCGAGAG 3193

RESULT 10
 US-10-450-763-24017
 ; Sequence 24017, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; PRIOR FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/549,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 24017
 ; LENGTH: 7856
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIMILAR
 ; LOCATION: (1697)..(4054)
 ; OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number
 ; OTHER INFORMATION: X63546, Smith-Waterman Score=423.
 ; US-10-450-763-24017

Alignment Scores:
 Pred. No.: 1,62e-191 Length: 7856
 Score: 2191.50 Matches: 408
 Percent Similarity: 88.2% Conservative: 33
 Best Local Similarity: 81.6% Mismatches: 58
 Query Match: 73.5% Indels: 1
 Db: 9 Gaps: 1

US-10-071-838-2 (1-549) x US-10-450-763-24017 (1-7856)

QY 1 MetAspValAlaGluValAlaGlySerTrpAlaGlnGluArgGluAspPheIleMet 20
 Db 1697 ATGACATGGTATAGAGATCAATAGTTTTCAGGCAAGGAGCGGAAAGCACTACTATAG 1756

QY 21 LysTYGluYseglYHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 Db 2834 ACCCTGCAAGGGGTATAGAGGGCCCTCCAGGCCACCAAGCCCAATTCAGAGGCGCC 2893

Db 1757 AAGTATGACAGAGGACACCGAGCTGGCTCCAGAGGACCAAGGGGCTGAGCCGTT--- 1813
 QY 41 SerTYAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
 Db 1814 GGATCAACAGCAGCATATGATCTTTTGGCATTTTGCATATGACGAGAGCTGCTCTGTG 1873
 QY 61 ThrAlaArgGluAlaYseglYleuArgGluIleSerArgLysSerLysTrpValAsp 80
 Db 1874 ACTGCACGGGAGCGGAGAAATTCGGCGGAGATGACAGACAGCAAGTGGATGAA 1933
 QY 81 MetLeuGlyAspTrpGluYseglYLysSerSerArgLysLeuIleAspArgAlaTYlys 100
 Db 1934 ATGCTGGAGAAATGGAGACATATTAAGCAGATGAACTCATATGATGAGTGTCAAG 1993
 QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
 Db 1994 GGAATTCATGAAACATCCGGGGCCGGGTGTGATGCTCTCTGAAACATTCAGAAATC 2053
 QY 121 LysLeuLysAsnProGlyArgTYGlnIleMetLysGluYseglYLysArgSerSerGlu 140
 Db 2054 AAGTTAAAAACCCCGAAGATTCAGATCATTAAGAGAGGGGCAAGGTCATCTGAA 2113
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 2114 CACATCACCAATGACCTGAGCGTGAAGAGAGACTCTCCGAAACATGCTCTTTAGG 2173
 QY 161 AspArgTYGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYGlnGluTYr 180
 Db 2174 GATCGATATGAGCCCAAGGAGGAGCAATTCATCATCTCTCTGAGCTATTCGAGTAT 2233
 QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTYr 200
 Db 2234 AACCCGAGGTGGGCTACTGAGGAGCTGAGCAGATACCGCTTGTTCCTTTAT 2293
 QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 2294 CTGCGTGAAGAGAGGATTCGGGACCTGGTGGAGCTGCTGGCAGTGAAGGAGCTCC 2353
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 Db 2354 CTGCCAGATTCACAGCCCAATGTGTGGAGACGTCCAGGGGCTCCAAAGCAACAGAG 2413
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 2414 CATGTGATCCCAAGTCAACCAAGACATGTGGCATCAAGCAAGGAGTCTATTC 2473
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 2474 GGGCAGTGTGCTCGTTAGGCTGCTCTCCGGAACCTGATTTGACGGATCTCTCGGG 2533
 QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGlnGlyGlnAlaLeuMetProIle 300
 Db 2534 CTCACCTGCGGCTGTGGAGCTGATTTGGTGAAGGAAAGAGTGTGATGCCAATA 2593
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysValGluThrLysThrSerArgCysGlyPro 320
 Db 2594 ACCAGCATTCCTTAAGGTTCAGAGAGCCCTCATGAAACATCCAGGTGTGGCTGTG 2653
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 2654 TGGGACGTCTGGGAGCAACATTTCTGATCTGGGCAATGACAGATGACCTGTCTC 2713
 QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
 Db 2714 AAGCATCTTAAGGCTCTTACGAAACATTAACAAGGAGGAGGAGGAGGAGGAGGAG 2773
 QY 361 AAlayProglunglInglYserSerAlaSerArgProValProAlaSerArglYglYlys 380
 Db 2774 GCCAAACGCGAGCAAGGGCTCTTGCCACCCAGGCTGTGCGGCTTACCGTGGGAG 2833

QY 381 ThleuCyelyseglYasparGlnAlaProProglYProProAlaAargPheProArgPro 400
 Db 2834 ACCCTGCAAGGGGTATAGAGGGCCCTCCAGGCCACCAAGCCCAATTCAGAGGCGCC 2893


```
Qy 401 ILETRSERALASERPROPROAAGLAProAArgSerSerThrProCySerProGlyValAla 420
Db 2894 ATTTGCTCAGCTTCCCGCCGATGGGCACTCTGTTTTCACGCCCTTCTCGGGGGGCT 2953
Qy 421 ValArgGluAerThrTyProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 2954 GTCCGGGAAGACAGTACCTCTGGGCACTCAGGGTGTGCCAGCTGGCCCTGGCTGCAG 3013
Qy 441 GlyGlyProGlnGlySerThrPArgPheLeuGlnTrpAserMetProAArgLeuProThr 460
Db 3014 GAGAGACCTCAGGGGTTCTCGAGATCTCTGGAGTGAAGTCAATGCCCGGCTCCCAAG 3073
Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyTrpAerPheArgGlnSerCysTrpVal 480
Db 3074 GACCTGGATATAGGGGGCCCTTGCTCCCATTTATATTTTGAACGAGCTGCTGGGCTC 3133
Qy 481 ArgAlaIleSerGlnGluAerGlnLeuAlaProCysTrpGlnAlaGlnIleProAlaGln 500
Db 3134 CGTGCTATATCCACGAGAGACAGCTGCCACTGCTGGACGCTGAACACTGGGAGAG 3193

RESULT 11
US-09-962-436-562
; Sequence 562, Application us/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppel, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689230-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 562
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-562

Alignment Scores:
Pred. No.: 1,62e-191 Length: 7878
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 3 Gaps: 1

US-10-071-838-2 (1-549) x US-09-962-436-562 (1-7878)
Qy 1 MetAspValValGluValAlaGlySerTrpAlaGlnGluAerGluAerIleMet 20
Db 1697 ATGACATGTGTAGAAATGCAATAGTTTGCAGGCAAGAGCGGAAGACATCTATG 1756
Qy 21 LysTyrgLulysGlyHisArgAlaGlyLeuProGluAerGlyGlyProLysProPheArg 40
Db 1757 AAGATATGACAAAGGACACGAGCTGGCTGCCAAGGACAAAGGGCCCTGAGCCCGTT-- 1813
Qy 41 SerTyraAsnAsnValAspHisLeuGlyIleValHisGluTrpGluLeuProLeu 60
Db 1814 GGAATCAACACAGCATTCGTTTGTGCAATTTTGCATGAGACGAGACCTGCTCTGTG 1873
Qy 61 ThrAlaArgGluAlaValGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGACGCGAGCGCAAGAAATTCGCGGAGATGACACGAACGAGAGTGGATGGA 1933
Qy 81 MetLeuGlyAerTrpGluLysTyrlsSerSerArgLysLeuIleAspArgAlaTyrls 100
Db 1934 ATGCTGGAGATGGAGACATATTAAGCACAGTACCAAACTCATGATCGAGTGTACAAG 1993
```

```
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnIleMet 120
Db 1994 GGAATTCCTCATGAACATCCGGGGCCGGGTGTGTCACTCTCTGTAACATTCAGGAATC 2053
Qy 121 LysLeuLysAsnProGlyValArgTyrlsIleMetLysGluLysGlyValArgSerSerGlu 140
Db 2054 AAGTTGAAAACCCCGGAAGATACAGATCATAGAGAGAGGGGCAAGAGTCAATCGAA 2113
Qy 141 HisIleGlnArgIleAspArgPheValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCACTGACCTGACGTGAGAGACATCTCCGGAACCATGCTTTTAGG 2173
Qy 161 AspArgTyrlsTyrlsGlnArgGluLeuLeuHisIleLeuLeuAlaTyrgLulysTyrls 180
Db 2174 GATCGATATGAGACCAAGACAGAGAGAACTATTCTACATCTCTCGGCTATTGGAGTAT 2233
Qy 181 AsnProGluValGlyTyrcysArgAerLysSerHisIleAlaAlaLeuPheLeuTyrls 200
Db 2234 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCAACCCCTTGTTCTCTTTAT 2293
Qy 201 LeuProGluGluAsnAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluAerGHisSer 220
Db 2294 CTGCTGAGAGAGACGCAATTCGAGCACTGGTGCAGCTGCTGACGTGAGAGGCACTCC 2253
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAerGlnGln 240
Db 2354 CTGCCAGATTCCTCAACAGCCCAATGTGGGACAGTCCAGGGGCTCCCAAGCAACAGAG 2413
Qy 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAerLysValAspLeuCys 260
Db 2414 CATGTGGTATCCCAAGTACAAACCAACCAATGTGGATCAGACAGAAAGGTCTATATGC 2473
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTCTCTGTAGGCTCTCTCTCGGAACCTGATGAGGGAATCTCTTCGGG 2533
Qy 281 LeuThrLeuArgLeuTrpAerValTyrlsLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 2534 CTCAACCTCGGCTCTGTGGACGTGTATTTGTGGAGAGAAACAGGTGTGATGCCAATA 2593
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGATTCCTCTTAAGTTTACAGAAACGCGCTCATGAAGACATCCAGGTGTGGCTG 2653
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAerThrValLeu 340
Db 2654 TGGCAGAGTCTGCGGAACCAATTCCTGATACCTGGGCAATGAACGATGACCCGTCTC 2713
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAerLeuProProPro 360
Db 2714 AAGCATTTTGGGCTCTTACAGAACTTAACAAGAAAGGAGGACCTGCCACCCCA 2773
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 2774 GCCAAAGCGCAGAGAGGTCTTGCGCACCGAGGCTCTGCGGCTTACGTTGTTGGGAAG 2833
Qy 381 ThrLeuLysGlyValAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 2834 ACCCTTCGCAAGGGGTATGAGCAGGCCCTCCAGGCCCAACAGCCAGTTCCAGGGGCC 2893
Qy 401 IleTrpSerAlaSerProProAerAlaProArgSerSerThrProCysProGlyValAla 420
Db 2894 ATTTGCTCAGCTTCCCGCCAGTGGGCACTCTGTTTTCACGCCCTTCTCGTGGGGCT 2953
Qy 421 ValArgGluAerThrTyProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 2954 GTCCGGGAAGACAGTACCTCTGGGCACTCAGGGTGTGCCAGCTGGCCCTGGCTGCAG 3013
Qy 441 GlyGlyProGlnGlySerThrPArgPheLeuGlnTrpAserMetProAArgLeuProThr 460
Db 3014 GAGAGACCTCAGGGGTTCTCGAGATCTCTGGAGTGAAGTCAATGCCCGGCTCCCAAG 3073
```


QY 461 AepLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgInsSerCysTyrVal 480
DB 3074 GACCTGGATATAGGGGCCCCCTTGTTCCCCCATTAAGTATGATTCAGAGGAGCTGCTGGTC 3133
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnHisProAlaGln 500
DB 3134 CCGGCATATCCAGAGGAGGACGAGCTGCGCACCCTGCGAGGCTGACACATCGGAGAG 3193

RESULT 12
US-10-843-641A-3021
; Sequence 3021, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3021
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3021

Alignment Scores:
Pred. No.: 1,62e-191 Length: 7878
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 9 Gaps: 1

US-10-071-838-2 (1-549) x US-10-843-641A-3021 (1-7878)

QY 1 MetAspValAlaGluValAlaGlySerTyrTPAlaGlnGluAspGlnSerIleIleMet 20
DB 1697 ATGCACATGCTGAGATTCAGATACATGTTGTCAGGACGACGAGACGAGACATACATATAG 1756
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 1757 AAGTATGACAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGGCTGAGCCCGTT--- 1813
QY 41 SerTyrAspAsnAsnValAspHisLeuGlyIleValHisGlyThrGlnLeuProProLeu 60
DB 1814 GGATTCACACAGCATATGATGCTTTGGCATTTTGCATGACAGACGAGCTGCTCTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerIleTyrValAsp 80
DB 1874 ACTGACCGGAGCGCAAGAAATTCGCGGAGATACACGAAACGACAAAGTGTATGAA 1933

QY 81 MetLeuGluAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 1934 ATGCTGGAGAAATGGAGACATATAGACAGTAGCAACATCATATGATGATGATCAAG 1993
QY 101 GlyMetProMetAsnIleArgLysProMetTyrSerValLeuLeuAsnIleGluGluMet 120
DB 1994 GGAATTCATGAAATCCGGGGGCGGTGTGTGATGCTCTCTGAAATTCAGGAAATC 2053
QY 121 LysLeuLysAsnProGluValArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 2054 AAGTTGAAAAACCCGGAAGATACAGATCAGAAAGGAGGCGCAAGGATCATCTGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrThrLeuArgLysHisIlePhePheArg 160
DB 2114 CACATTCACACCATGACCTGAGCGAGAGACGACTCTCGGAACCATGATCTTCTTAG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 2174 GATCGATATGAGCGCAGAGGAGAACTATTCATCTCTCTGCGCTTATCGAGAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
DB 2234 AACCCGAGGTGGGCTACTGACGAGACCTGACCATCACCGCCTTGTCTCTTAT 2293
QY 201 LeuProGluLysAlaPheThrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 2294 CTGCTGAGAGAGAGGACATTCGAGCAGTGTGAGCTGTGCGCAGTGAAGGACCTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 2354 CTGCCAGATTCACAGCCCAATGTGTGGAAGTCCAGGGGCTCCAAACCAACAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 2414 CATGTGTATCCCAAGTACCAACCAACCAATGTGTGATCAGAGACAAAGAGTATATC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
DB 2474 GGGCAGGTGGCTCGTGTAGGCTGCTTCCGGAACCTGATGAGCGGATCTCTCGGG 2533
QY 281 LeuThrLeuArgLeuThrAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB 2534 CTACACCTGCGGCTGTGGACGTGATTTGGTGAAGGAAACAGGTGTGATGCCAAT 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 2594 ACCAGCATTTGCTTAAGTTACAGAGAGGCGCTCATGAAACATCCAGGTGTGCGT 2653
QY 321 TPAlaArgPheCysAsnArgPheValAspThrTPAlaArgAspGluAspThrValLeu 340
DB 2654 TGGGACGCTGCGGAAACCAATCTTCATACCTGGGCGCATGACGATACACCGGCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 2714 AAGCATTTAGGGCTCTTACAGAAACTAACAGAGAGAGGAGGAGCTGCCACCCCA 2773
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgLysGlyLys 380
DB 2774 GCCAAACGAGACAAAGGCTCTTGGCACCCAGGCGCTGTGCGGCTTACGTGTGGAG 2833
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 2834 ACCCTCTGCAAGGGGTATAGGAGGAGCCCTCCAGGCGCCACCGCATTCACAGGGCCC 2893
QY 401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 2894 ATTGCTCAAGCTTCCCGGCAATGAGGCAATCTGTTTTCACGCCCTGCTGTGAGGCT 2953
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 2954 GTCCGGGAAGACAGTACCTGTGGCACTAGGGGTGTGCCAGGCTGCGCTGCTAG 3013

```
Qy 441 G1yG1yProGInglYserTPrpAphLeuGInTPraSerMetProArGLeuProThr 460
Db 3014 GGAGACCTCAGGAGGCTTCTGGAGATTCCTGGAGTGAAGTCAATGCCCGGCTCCACAG 3073
Qy 461 AspLeuAhpValG1uG1yProTPPheArgH1sTyArpPheArgGlnSerCyETrpVal 480
Db 3074 GACCTGGATATTAAGGGGGCCCTTGCTGCCCATATATGATTTTGAACGAGCTGCTGGGTC 3133
Qy 481 ArgAlaIleSerGIngluAhpGlnLeuAhpProCySerTPPInlaGluH1sProAlaGlu 500
Db 3134 CGTCCATATCCACAGAGGACCTGGCCACTGCTGGCAGGCTGAACACTGGCGAGAG 3193

RESULT 13
US-10-956-157-1259
; Sequence 1259, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mount, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1259
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1259

Alignment Scores:
Pred. No.: 1.62e-191 Length: 7878
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 9 Gaps: 1

US-10-071-838-2 (1-549) x US-10-956-157-1259 (1-7878)
Qy 1 MetAhpValG1uValAlaG1ySerTPPAlaGIngluArgGluAhpIleIleMet 20
Db 1697 ATGACATCTGTAGAGAAATGACATAGTTTGCAGGACAGAGCCGAGAGACATCTTATG 1756
Qy 21 LysTPGluYsg1yH1sArgAlaG1yLeuProGluAhpYsg1yProLYserProPheArg 40
Db 1757 AAGATATGACAGAGGACACCGAGCTGGCTGCCAGAGGACAGAGGGGCTGACCCGTT-- 1813
Qy 41 SerTPraSerAhpValAhpH1sLeuG1yIleValH1sGluTPHrGluLeuProLeu 60
Db 1814 GGAATCAACAGCAGCATTTGATCGTTTGGCATTTTGCATGTAGAGCGAGCTGCTCTGTG 1873
Qy 61 ThrAlaArgGluAlaYsgInlIleArgArgGluIleSerArgLYserLYserTPValAhp 80
Db 1874 ACTGCACGGAGGGGAGAGAAATTCGCGGAGATGACACAGACGAAATGATGGA 1933
Qy 81 MetLeuG1yAhpTPGluYsg1yTyLYserSerArgLYserLYserAhpAlaTYLYs 100
Db 1934 ATGCTGGGAGAAATGGAGACATATATAGCAGTAGCAATCATATGATCGAGTATCAAG 1993
Qy 101 G1yMetProMetAhpIleArgG1yProMetTPPserValIleuLeuAhpIleGluGluMet 120
Db 1994 GGAATTCCTCATGACATCCGGGGCCGGGTGTGTCAGTCTTCGMAACATTCAGAAATC 2053
Qy 121 LysLeuLYserAhpProG1yArgTYsgInlIleMetLYsg1yG1yLYserSerSerGlu 140
Db 2054 AAGTTGAAAACCCCGGAGATACCAATCATTAAGAGAGAGGGGAGAGGTCTCATCTGA 2113
Qy 141 H1sIleGlnArgIleAhpArgAhpValSerG1yThrIleuArgLYsH1sIlePhePheArg 160
Db 2114 CACATTCACCACTGACCTGAGCTGAGTGAAGACACTCTCCGGAACCATGTCTTCTTAGG 2173
```

```
Qy 161 AspArgTYsg1yThrLYsgInaArgGluLeuH1sIleLeuLeuAlaTYrgIngluTYr 180
Db 2174 GATCGATATAGAGCCAAAGAGAGGAACTATTCTACATCTCTCTGGCCTATTGCGAGAT 2233
Qy 181 AhpProGluValG1yTYrCYsAhpAhpLeuSerH1sIleAlaIleuPheLeuTYr 200
Db 2234 AACCCGAGAGGTGGCTACTGACAGGACCTGAGCAACATCCACCTGTTGCTCTTAT 2293
Qy 201 LeuProGluGluAhpAlaPheTPAlaLeuValG1yLeuLeuAlaSerGluAhpH1sSer 220
Db 2294 CTGCTGAGGAGAGAGCATTTCTGGCATGTGTCAAGCTGCTGGCCAGTGAAGACTCC 2253
Qy 221 LeuGInglYpHeH1sSerProAhpG1yTYrThValGlnG1yLeuGlnAhpGInglInglu 240
Db 2354 CTGCAAGATTCACAGGCCAAATGAGGAGCATCCAGGGGCTCCAAAGCAACAGAG 2413
Qy 241 H1sValValAlaThrSerGlnProLYserTPHrMetG1yH1sGlnAhpLYsAhpLeuCYs 260
Db 2414 CATGTGGTACCCAAAGTCAACCAAGACCATGTGGCATTCAGACACAGAAAGGTCTATGC 2473
Qy 261 G1yG1yCySerProLeuG1yCyLeuIleArgIleLeuIleAhpG1yIleSerLeuG1y 280
Db 2474 GGGCAGTGTCTCTGTAGGCTGCTCTCGGAACTGATTAACGAGATCTCTCTGGG 2533
Qy 281 LeuThrLeuArgLeuTPraAhpValTYrLeuValG1uG1yGluGlnAlaLeuMetProIle 300
Db 2534 CTCACCTGCGCCTGTGGAGCGTATTTGTGGAGAGAAACAGGCTTGATGCCAATA 2593
Qy 301 ThrArgIleAlaPheLYsValGlnGlnLYsArgLeuThrLYsThrSerArgCYsgLYPro 320
Db 2594 ACCAGATTCCTCTTAAGGTTCAGAGAGCGCTCATGAAAGCATCCAGGTGTGGCTG 2653
Qy 321 TPraAlaArgPheCYsAhpAhpPheValAhpThrTPraAlaArgAhpGluAhpThrValLeu 340
Db 2654 TGGCAGCTGTGGGAAACCAATTCCTGATACCTGGGCCATGAAACGATGACACCGTCT 2713
Qy 341 LYsH1sLeuAhpArgAlaSerMetLYsLYsLeuThrArgLYsgInglYsAhpLeuProPro 360
Db 2714 AAGCATTTTGGGCTCTTACGAAAGAACTAACAGAGAGAGAGGAGCTGCCACCCCA 2773
Qy 361 AlaLYsProGluGlnG1ySerSerAlaSerArgProValProAlaSerArgG1yG1yLYs 380
Db 2774 GCCAAACGCGAGCAAGGGTCTTGGCACCCAGGCTGTGGCTTCACTGTGTGGAG 2833
Qy 381 ThrLeuCYsLYsG1yAhpArgGlnAlaProProG1yProProAlaArgPheProArgPro 400
Db 2834 ACCCTTGCAGAGGATATAGCAGGCCCCCTCCAGGCCACACAGCCAGATTCACGGGCC 2893
Qy 401 IleTPraSerAlaSerProProArgAlaProArgSerSerThrProCYsProG1yG1yAla 420
Db 2894 ATTTCCTACCTTCCCGGCATAGGCAATCGTTTTCACGCGCTCTCTGTGGGGCT 2953
Qy 421 ValArgGluAhpThrTYrProValG1yThrGlnG1yValProSerProAlaLeuAlaGln 440
Db 2954 GTCCGGGAGAACAGTACCTGTGGGCACTCAGAGGTGTGCCAGCGCTGGCTGCAG 3013
Qy 441 G1yG1yProGlnG1ySerTPraArgPheLeuGlnTPraSerMetProArgLeuProThr 460
Db 3014 GAGAGACCTCAGGATTCCTGAGATTCCTGAGTGAAGTCAATGCCCGGCTCCAAAG 3073
Qy 461 AspLeuAhpValG1uG1yProTPPheArgH1sTyArpPheArgGlnSerCyETrpVal 480
Db 3074 GACCTGGATATTAAGGGGGCCCTTGCTGCCCATATATGATTTTGAACGAGCTGCTGGGTC 3133
Qy 481 ArgAlaIleSerGIngluAhpGlnLeuAhpProCySerTPPInlaGluH1sProAlaGlu 500
Db 3134 CGTCCATATCCACAGAGGACCTGGCCACTGCTGGCAGGCTGAACACTGGCGAGAG 3193

RESULT 14
US-10-450-763-24016
; Sequence 24016, Application US/10450763
; Publication No. US20050196754A1
```

GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450, 763
PRIOR FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 24016
LENGTH: 8180
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (4369) (6126)
OTHER INFORMATION: 97% homologous to Homo sapiens oncogene, accession number
US-10-450-763-24016
OTHER INFORMATION: X63546, Smith-Waterman Score=3063.
Alignment Scores:
Pred. No.: 3, 88e-186 Length: 8180
Score: 2133.50 Matches: 408
Percent Similarity: 77.2% Conservative: 33
Best Local Similarity: 71.5% Mismatches: 58
Query Match: 9 Indels: 72
DB: 9 Gaps: 2
US-10-071-838-2 (1-549) x US-10-450-763-24016 (1-8180)
QY 1 Metaspvialvalgluvalaaglysertrpalaingluargluabpillemet 20
DB 1697 ATGACATGCTGTAAGATGACATGATGTTGACAGCCAGAGCGGAAGACATCTTAG 1756
QY 21 Lygylrgluylgylhialrgalaglyleuproglyuabpilygylprolysprophearg 40
DB 1757 AAGTATGACAGAGGACACCGAGCTGGCTGCAGAGACAGAGGCGCTGAGCCCTT--- 1813
QY 41 SerTyraasnaasnaValasphlsleuglyllevalHiegluThrgluLeupropoleu 60
DB 1814 GGAATCAACAGACGATGATGCTTTGGCATTTGACATGACGAGCTGCTCTGTG 1873
QY 61 ThrAlaargluAlaIyseglnlleargatrgluileserarglysertrpValaap 80
DB 1874 ACTGACGGGAGGCGGAGAAATTCGGCGGAGATGACAGAACGACAGAGTGTGAGAA 1933
QY 81 MetleuglyAserTrpGluylrlyserSerarglyleuileaphrgalarglylys 100
DB 1934 ATGCTGGGAGATGGGACATATTAAGACAGTACCAACTCATAGATGATGATCAAG 1993
QY 101 GlymetPromerAenlleargglyPrometTrpSerValleuileuamilleglumet 120
DB 1994 GGAATTCATCAACATTCGGGACCGGCTGCTGCTCTCTGAACTTCAAGAAATC 2053
QY 121 LygbleuylAenProglYargTrglnilemetlysgluyllysaargserSerlu 140
DB 2054 AAGTGAAGAAACCCCGGAGATACCATCATGAGAGAGGCGGAGAGGCTCATTTGAA 2113
QY 141 HislleglinaargyleaphrgaapValserglyThrleuarglyHisllephepharg 160
DB 2114 CACATTCACACATGACCTGACGCGAGAGAGACGACTCTCGGAAACCATGTCTTTAGG 2173
QY 161 AsparglyTrgylThrlyeglnarggluleuileuHislleleuileuAlayrgluGlyTr 180
DB 2174 GATTCATATGAGCCAGAGAGGAACTATTCTATCTCTCTGCGCTTATTCGAGTAT 2233
QY 181 AenProgluValglYTrCyargarbLeuSerHislleAlaaleuPheleuLeuTr 200

DB 2234 AACCGGAGGTGGGCTACTGACAGGACCTGAGCCATCATCAGCGCTTGTCTCTTAT 2293
QY 201 LeuProgluGluabpAlaPheTrpAlaLeuValGlnleuileuAlaSerGluargHiser 220
DB 2294 CTGCTGAGAGAGACCATCTCGGACCTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCT 2353
QY 221 LeuGlnGlyPheHiserProaenGlyThrValGlnGlyLeuGlnleuPheGlnGln 240
DB 2354 CTGCGGATTCACAGCCCAATGAGGACGTCCAGGGGCTCCAAAGCCAAACAGAG 2413
QY 241 HisValValAlaThrSerGlnProlyThrMetGlyHisGlnaaplylysaPheLeuCy 260
DB 2414 CATGTGTACCAAGTACCAACCAAGACATGTCATCAGGACAGAGAGCTATGC 2473
QY 261 GlyGlnCySerProleuGlyCybleuileargyleuileaphrglylleserleuGly 280
DB 2474 GGGCAGTGTGCTCTGTTAGGCTGCTCTCGGACCTGATGACCGGATCTCTCGGG 2533
QY 281 LeuThrleuArgleuTrpAerValTyrlleuValGlnGlyGlnAlaLeuMetProile 300
DB 2534 CTGACCTTGCTGCTGAGCGTGTATTTGTGGAAGAGAACAGGTGTGTATGACCAATA 2593
QY 301 ThrArglleAlaPheleValGlnGlnlysaTrgLeuThrlyThrSerargCyGlyPro 320
DB 2594 ACCAGATTCCTCTTAAGTTTACAGAGAGCGCTCATAGAACATCCAGGTGTGCTG 2653
QY 321 TrpAlaargPheCyasnaargPheValAerThrTrpAlaargapgluabpThrValleu 340
DB 2654 TGGGACGCTGCTCGGAAACCAATTCCTGATCTGAGGACATGAACATGACACCGTCTC 2713
QY 341 LygHisleuArgAlaSerMetlylyleuThrArglyGlnGlylysaPheLeupropoleu 360
DB 2714 AAGCATTTAGGCTCTTCAAGAACTTAACAGAGAACAGAGGACCTGCAACCCCA 2773
QY 360 ----- 360
DB 2774 GGCCCAACAGCCCTGGGAGCAAGGTGTGTGAGAGAGGCCCAAGCCACTGTGAACCTG 2833
QY 360 ----- 360
DB 2834 GGGGAGTCCAGAGACCAACCATGCCCCAGCGCTTCCCATGCGAGGACAGACAC 2893
QY 360 ----- 360
DB 2894 ACCCTCCTCTGAGATGACAGACTACAGCGGTGTCTGATGTACAGCAACAGGGGCTC 2953
QY 361 ----- 361
DB 2954 AACAGAGACCCCAAGAGCTCCAGAGATGACACCAAGCCGAGAGAGGTCTTGGACAC 3013
QY 370 rArgProValProAlaSerargGlyGlylysaThrleuCylysglyAsparGlnAla 390
DB 3014 CAGGCTGTGCGGGCTTACGTTGTGGAGAACCTCTGCAAGGGGTATAGGACAGGCCCC 3073
QY 390 oProglYProProAlaargPheProargProileTrpSerAlaSerProProargAla 410
DB 3074 TCCAGGCTCCAGGACCAAGTTCAGCGGCTTGTCTAGCTTCCCGCCCATGGGCAATC 3133
QY 410 oArgSerSerThrProCyserProglYglValaValaArgGluAerThrTyPrProValGlyTh 430
DB 3134 TCGTTTTTCCAGGCTCTGCTGCTGAGGCTCTCCGGAGAGACGTAACCTGTGGGAC 3193
QY 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPhele 450
DB 3194 TCAGGGTGTGCCCAAGCTGAGCTGCTGAGGAGAGACCTGAGGCTTCTGAGATTCCT 3253
QY 450 uGlnTrpAenSerMetProargleuProThrAspLeuAerValGlnGlyProTrpPhe 470
DB 3254 GAGTGAAGATTAAGTCCCGGCTCCCAAGGACCTGATTAAGAGGGGCTTGTGTCCC 3313
QY 470 gHisTyraapPheargGlnSerCyserTrpValaArgAlaIleSerGlnGlnleuAla 490
DB 3314 CCAATTATATTTTGAACGAGCTGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3373

QY 490 aProCy5TGTGlnAlaGlnHisProAlaGln 500
DB 3374 CACCTGCTGGCAGGCTGAACACTGGCGAGAG 3404

RESULT 15
US-10-283-975A-272
Sequence 272, Application US/10283975A
Publication No. US20040110792A1
GENERAL INFORMATION:
APPLICANT: Ortho-Clinical Diagnostics, Inc.
TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
FILE REFERENCE: CDS 293 PCT
CURRENT APPLICATION NUMBER: US/10/283, 975A
PRIORITY FILING DATE: 2002-10-30
PRIORITY APPLICATION NUMBER: 60/340, 938
PRIORITY FILING DATE: 2001-10-30
PRIORITY APPLICATION NUMBER: 60/338, 997
PRIORITY FILING DATE: 2001-10-30
PRIORITY APPLICATION NUMBER: 60/340, 081
PRIORITY FILING DATE: 2001-10-30
PRIORITY APPLICATION NUMBER: 60/341, 012
PRIORITY FILING DATE: 2001-10-30
NUMBER OF SEQ ID NOS: 900
SOFTWARE: PatentIn version 3.1
SEQ ID NO 272
LENGTH: 8201
TYPE: DNA
ORGANISM: HUMAN
US-10-283-975A-272

Alignment Scores:
Pred. No.: 3.9e-186 Length: 8201
Score: 2133.50 Matches: 408
Percent Similarity: 77.2% Conservative: 33
Best Local Similarity: 71.5% Mismatches: 58
Query Match: 71.5% Indels: 72
Gaps: 2

US-10-071-838-2 (1-549) x US-10-283-975A-272 (1-8201)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 1697 ATGCAACATGTTGATGATGCAATGATGTTGACAGCACAAGCGGAAAGACATCTATG 1756

QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 1757 AAGTATGACAAAGGAGACACCGAGCTGGCTGCCAGAGGCAAGGGGCTGAGCCCGTT-- 1813

QY 41 SerTyrTrpMetAsnValAspHisLeuGlyIleValHisGluTrpGluLeuProLeu 60
DB 1814 GGAATCAACAGCAGCATTTGATCGTTTGGCATTTTGCATGAGACGAGCTGCTCTGTG 1873

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgLuiLeseArgLysSerLysTrpValAsp 80
DB 1874 ACTCCAGCGGAGCGGAGAAATTCGGCGGAGATGACACGAAAGAGCAAGTGTGAA 1933

QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 1934 ATCTGGGAGAAATGCGAGACATATAGCACAAGTACAACTCATAGTGTACAG 1993

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 1994 GGAATTTCCCATGAACATCCGGGGCCCGGTGTGTAGTCTCTCGAACATTCAGGAATTC 2053

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 2054 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAGAGCGGCAAGAGTCAATCGAA 2113

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 2114 CACATCCACACACATCGACTGAGCGTGAAGAGAGACTCTCCGGAACCATGTCTTTTAGG 2173

QY 161 AspArgTyrGlyThrLysGlnArgGluLeuHisIleLeuLeuAlaTyrGluGlnTyr 180
DB 2174 GATCGATATGAGCCCAAGAGAGGAACTATTATCATCTCTCTGGCTATTGAGATAT 2233

QY 181 AsnProGluValGlyTyrCyatArgLeuSerHisIleAlaIleLeuPheLeuTyr 200
DB 2234 AACCCGAGGTGGCTGACTGACAGGACCTGAGCCACATCAACCGCTGTGCTCTTTAT 2293

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 2294 CTCTCGAGAGAGACGATTTCTGGCATCTGCTGCACTGCTGGCCAGTGAAGGACCTCC 2353

QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnArgGlnGln 240
DB 2354 CTGCCAGGATTCACAGCCCAATATGGTGGGACAGTCCAGGGCTCCCAAGCAACAGAG 2413

QY 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAspLysValAspLeuCys 260
DB 2414 CATGTGGTACCCCAAGTCACAAACCAAGACCATGTGGCATTCAGACAAAGAAAGGTCTATGC 2473

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 2474 GGGCAGGTGCTCTGTTAGGCTGCTCTCTCGGAACCTGATTCAGGATCTCTCTGGG 2533

QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
DB 2534 CTCACCTGCTGGCTGTGGGACGATGATTTGGTGAAGAGAAACAGGATGTTGACCAATA 2593

QY 301 ThrArgIleAlaPheLysValGlnGlnLysValGlyLeuThrLysThrSerArgCysGlyPro 320
DB 2594 ACCAGCATGCTCTTAAAGTTGACAGAAACCGCTCATGAAACATTCACAGTGTGGCTG 2653

QY 321 TrpAlaArgPheCysAsnArgPheValAspTrpTrpAlaArgAspGluAspThrValLeu 340
DB 2654 TGGCAGGTGCTGGGAAACCAATTTCTGATACCTGGGCCATGAACATGACACCTGCTC 2713

QY 341 LysHisLeuLysArgLaseMetLysLysLeuThrArgLysGlnIleAspLeuProProPro 360
DB 2714 AACATCTTAAAGGCTCTTACGAAAGAACTAAAGAAAGCAAGGAGGACCTGCCACCCCA 2773

QY 360 ----- 360

QY 2774 GGCCCAACAGCCCTGGGACGAAAGTGTGTGGCAGAAACCCAGCCAGTCTGAACCTTG 2833

QY 360 ----- 360

QY 2834 GGGGAGTCCCAAGAGCCACCCACATGCCCAAGGCTTCCCATGCCAGGACGACAC 2893

QY 360 ----- 360

QY 2894 ACCCTCTCCTCTGGGATCAGCAGACTACAGGCGTGTGTCAGTGTCAACACAGGGGCC 2953

QY 361 -----AlaLysProGluGlnGlySerSerAla 370

QY 2954 ACAAGAGAACCCCAAGGACTCCAGAGTGCAGGCAACGAGCAAGAGGTCCTTGGAC 3013

QY 370 ArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyLysAspArgGlnAlaPr 390

QY 3014 CAGGCTGTGCTCGCTTCACTGTGTGGAAAGACCTCTGCAAGGGGATATAGGACGCGCC 3073

QY 390 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPr 410

QY 3074 TCCAGGCGCCACAGCCAGTTCCAGGGGCCCATTTGCTCAGCTTCCCGCCATGGGCTATC 3133

QY 410 ArgSerSerThrProCysProGlyGlyValAlaValArgGluAspThrTyrProValGlyTh 430

QY 3134 TCGTTTTTCCAGCGCTCTGCTGTGGGGCTGTCGGGAAAGACACGTAACCTGTGGGAC 3193

QY 430 ArgGlnValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLe 450

QY 3194 TCAAGGAGTCCCAAGCTGCGCTGAGGAGGACCTCAAGGATTCCTGAGAAATTCCT 3253

QY 450 uGIntTrpAsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheAr 470

Thu Apr 6 16:21:15 2006

Db	3254	GAAGTGAAGTCAATGCCCGGCTCCCAAGGAACTGATATATAGGGGGCCCTTGTTCCC	3313
Qy	470	GHISLYRAAPPheargGlnSerCySTRpValArgAlaIleSerGlnGlnuapGlnLeuAl	490
Db	3314	CCATTATGATTGAAAGGAGCTGCTGGGTCCGTGCCATATCCAGAGAGACCAAGCTGGC	3373
Qy	490	aProCySTRpGlnAlaGlnHisProAlaGlu	500
Db	3374	CACCTGCTGGCAGAGCTGAACACTGCGGAGAG	3404

Search completed: April 5, 2006, 15:15:55
Job time : 1256 secs

LINE 40005 R1C10 16:21:15

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:49:32 ; Search time 289 Seconds

(without alignments)
3376.752 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983
Sequence: 1 MDVVEVAGSMWAGREDIIM.....TSGRCICGLHSSQPPRGR 549

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ .p2n.model -DEV=xlh
-Q=/abes/ABSSWB.spool/US10071838/runcat.05042006.141917.5227/app.query.fasta_1
-DB=Issued Patents NA -OPMT=fastaop -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abes07 -USBR=US10071838 -QCEN 1 1 193 -runcat.05042006.141917.5227 -NCPU=6
-ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/1/COMB.seq.*
2: /cgn2_6/prodata/1/ina/5/COMB.seq.*
3: /cgn2_6/prodata/1/ina/6/COMB.seq.*
4: /cgn2_6/prodata/1/ina/6/COMB.seq.*
5: /cgn2_6/prodata/1/ina/6/COMB.seq.*
6: /cgn2_6/prodata/1/ina/6/COMB.seq.*
7: /cgn2_6/prodata/1/ina/6/COMB.seq.*
8: /cgn2_6/prodata/1/ina/6/COMB.seq.*
9: /cgn2_6/prodata/1/ina/6/COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2850	95.5	2084	3	US-09-799-451-692
2	2827	94.8	2072	3	US-09-799-451-691
3	2133.5	71.5	8201	2	US-08-253-155A-9
4	344	11.5	1981	3	US-09-620-312D-715
5	337	11.3	1823	3	US-09-774-528-336
6	337	11.3	1823	3	US-10-120-988-336
7	332.5	11.1	1925	3	US-09-949-016-2139
8	258	8.6	3039	3	US-09-620-312D-675
9	253	8.5	2955	3	US-09-620-312D-676

10	238	8.0	1083	3	US-09-248-796A-5457	Sequence 5457, Ap
11	226	7.6	2456	3	US-09-949-016-1620	Sequence 1620, Ap
12	222	7.4	1887	3	US-09-270-767-12462	Sequence 12462, A
13	194.5	6.5	806	3	US-09-270-767-9899	Sequence 9899, Ap
14	182.5	6.1	4039	2	US-08-363-300-1	Sequence 1, Appl1
15	175.5	5.9	3983	3	US-09-762-111-3	Sequence 3, Appl1
16	175.5	5.9	3988	3	US-09-762-111-4	Sequence 4, Appl1
17	172	5.8	716	3	US-09-533-559-4537	Sequence 4537, Ap
18	172	5.8	1056	3	US-09-248-796A-1133	Sequence 1133, Ap
19	170.5	5.7	2574	3	US-10-104-047-400	Sequence 400, App
20	169	5.7	2066	3	US-10-104-047-56	Sequence 56, Appl
21	164	5.5	2013	3	US-09-502-540-1173	Sequence 1173, Ap
22	163.5	5.5	58857	3	US-09-477-962-1	Sequence 1, Appl1
23	162	5.4	61158	3	US-09-949-016-15041	Sequence 15041, A
24	162	5.4	75431	3	US-09-949-016-15122	Sequence 15122, A
25	158.5	5.3	9818	3	US-09-902-540-987	Sequence 987, App
26	156	5.2	17315	3	US-09-902-540-1103	Sequence 1103, Ap
27	155	5.2	30783	3	US-09-902-540-1258	Sequence 1258, Ap
28	154	5.2	1659	3	US-09-248-796A-1820	Sequence 1820, Ap
29	151.5	5.1	5418	3	US-09-949-016-2477	Sequence 2477, Ap
30	150.5	5.0	1902	3	US-09-902-540-8758	Sequence 8758, Ap
31	150.5	5.0	9556	3	US-09-902-540-929	Sequence 929, App
32	149	5.0	3073	3	US-09-902-540-643	Sequence 643, App
33	149	5.0	7527	3	US-09-252-991A-71	Sequence 71, Appl
34	148.5	5.0	1708	3	US-10-104-047-1901	Sequence 1901, Ap
35	148.5	5.0	2502	3	US-10-190-967-8	Sequence 8, Appl1
36	148.5	5.0	2631	3	US-09-902-540-4168	Sequence 4168, Ap
37	148.5	5.0	3763	3	US-09-919-039-243	Sequence 243, App
38	148.5	5.0	5185	3	US-09-976-594-640	Sequence 640, App
39	148.5	5.0	18686	3	US-09-902-540-1206	Sequence 1206, Ap
40	148.5	5.0	20757	3	US-09-902-540-1189	Sequence 1189, Ap
41	148.5	5.0	50780	3	US-09-902-540-1243	Sequence 1243, Ap
42	148	5.0	4800	3	US-09-902-540-562	Sequence 562, App
43	148	5.0	26012	3	US-09-902-540-1212	Sequence 1212, Ap
44	148	5.0	29272	3	US-09-902-540-1217	Sequence 1217, Ap
45	147.5	4.9	723	3	US-09-270-767-1491	Sequence 1491, Ap

ALIGNMENTS

RESULT 1
US-09-799-451-692
Sequence 692, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Aundt, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungting
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-Hong
APPLICANT: Wang, Zhilwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Keena
APPLICANT: Dimanac, Radoje T.
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163)..(1755)
US-09-799-451-692

Alignment Scores:

Pred. No.:	5,036-183	Length:	2084
Score:	2850.00	Matches:	528
Percent Similarity:	96.4%	Conservative:	1
Best Local Similarity:	96.2%	Mismatches:	2
Query Match:	95.5%	Indels:	18
	3	Gaps:	1

US-10-071-838-2 (1-549) x US-09-799-451-692 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 163 ATGGACGTGTAGAGTGTGGCGGCGGTGTGGGCACAAGACGAGACATCATTTATG 222
QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 223 AAATACGAAAGGAGACACCGAGCTGGCTGCCAGAGACAGAGGGCTTAAGCTTTTGA 282
QY 41 SerTyraAsnaAsnaValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
DB 283 AGCTACACAACAACGTGCATCTTTGGGATTTGATGACAGACGAGCTGCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 343 ACTGGCGGAGGCGAGACCAATTCGCGGAGATCACCCGAAAGACAGTGGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTrpHis 100
DB 403 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAAAGCTCATGATCGACGTAACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 463 GGAATGCCCATGAATCCGGGGCCGATGTGTGATCTCTCTGATACATGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerLys 140
DB 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 529 AACATCCAGCGCATGACCGGAGCGTAAGCGGACATTAAGAGACATTAATCTTCAGG 588
QY 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
DB 589 GATCGATACGGAAACCAAGCAGCGGGAATCTCCATCTCTGTCATATGAGAGAT 648
QY 181 AsnProGluValAlaGlyTrpCysArgAspLeuSerHisIleAlaIlePheLeuLeuTrp 200
DB 649 AACCCGGAGGTGGGTACTGACGAGGACCTGAGCCATCGCGCTTCTCTCTCAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlyAspHisSer 220
DB 709 CTCTCTGAGAGATGATCTCTGGGCACTGGTGCAGCTGTGGCCAGTGAAGGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyIleThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 769 CTGCAGGGATTTTCAACAGCCAAATGGCGGACCGTCCAGGGGCTCCAGACCAACAGAG 828
QY 241 HisValValAlaIleThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 260
DB 829 CATGTGGTACCACTCAACACCAAGACCAATGGGCGATCAGGACAAAGAAATCTATGT 888
QY 261 GlnGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 889 GGGCAGTCTTCCCTTGGCTGCTCATTCGGATATGATGAGGGGATCTCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrlleuValGlnGluGlnAlaLeuMetProIle 300

DB 949 CTCACCTGCGCGCTGTGGACGTGTATCTGTAGAGCCGAAACAGCGGTGATCCCATTA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTrpSerArgCysGlyPro 320
DB 1009 ACAAGAAATCGCTTTAAAGTTACAGCAAGAGGCTCTCAAGAAAGCTCCAGGTGTGGCCG 1068
QY 321 TrpAlaArgPheCysAsnaArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1069 TGGCAGATTTTTCACACCGGTTCTGTGATACCTGGGCGAGGATGAGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1129 AAGCATCTTGGGCTCTATGAAAGAACTAACAAGAAAGCAGGGGAGACTCTCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
DB 1189 GCCAAACCCGAGAGAGGTGTGTGGCATCCAGGCTGTGGCTTCCAGGTTCAGTGGCGGAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1249 ACCCTTGCAAGAGGGGACAGGCAAGCCCTTCAGGCCCAACAGCCGGTTCCGCGGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1309 ATTGTGTAGCTTCCCGCCACGGGCACTCTGTTCTTCCACACCTGTCTGTGGGGCT 1368
QY 421 ValArgGluAspTrpTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1369 GTCCGGAGAGACACTTACCTGTGGCATCTCAGGATGAGGTGTGCCACCGGCTGGCTCAG 1428
QY 441 GlyLysProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1429 GAGAGACTCAAGGTTCTGTGAGATTCCTGAGTGAATCCAGTCCCGGCTTCCAAAG 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTrpAspPheArgGlnSerCysTrpVal 480
DB 1489 GACCTGACAGTGAAGGGGCCCTTGGTTCGGCATTAATTCACAGACAGAGCTGTGGGCTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1549 CGTGCATATCTCCAGAGACACAGCTGACCCCTGCTGGAGGCTGAACACCTTGGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1609 CGGTTGAGATGGGTTTGTGTGCTGCACCAAGCACTGATTCGACAGAGGCAACCCCTTCA 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1669 GCTAGGAGACGAACAGCGGTGTGCTCCACCTCAAGGCGCTTGCCTCGGCTCCACTTG 1728
QY 541 GlusSerGlnPheProProGlyPhe 549
DB 1729 GAAAGTTCTCAGTTCCTTCCAGGCTTC 1755

RESULT 2
US-09-799-451-691
Sequence 691, Application US/09799451
Patent No. 6783869
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aiding J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yundong
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Zhi-hong
APPLICANT: Wang, Zhiwei


```
/ Patent No. 5691147
/ GENERAL INFORMATION:
/ APPLICANT: Gyrus, Jeno
/ APPLICANT: Draetta, Giulio
/ TITLE OF INVENTION: CDK4 Binding Proteins
/ NUMBER OF SEQUENCES: 95
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 60 State Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII(text)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/253,155A
/ FILING DATE: 02-JUN-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Vincent, Matthew P.
/ REGISTRATION NUMBER: 36,709
/ REFERENCE/DOCKET NUMBER: MII-028
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 227-7400
/ TELEFAX: (617) 227-5941
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 8201 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-08-253-155A-9

Alignment Scores:
Pred. No.: 6,2e-134 Length: 8201
Score: 2133.50 Matches: 408
Percent Similarity: 77.2% Conservative: 33
Best Local Similarity: 71.5% Mismatches: 58
Query Match: 71.5% Indels: 72
DB: 2 Gaps: 2

US-10-071-838-2 (1-549) x US-08-253-155A-9 (1-8201)

QY 1 MetAapValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGGCATGTGTAGAGAAATGCAAGATGTGTGACGACAGCGGAGCAATATCTATG 1756
QY 21 LysTrpGluValGlyHisArgAlaGlyLeuProGluAspGlyGlyProLysProPheArg 40
Db 1757 AAGATATGCAAGGAGCACCGAGCTGGCTGCCAGAGCAAGGGGCTGAGCCCGTT-- 1813
QY 41 SerTrpAaaAaaAaaValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db 1814 GGAATCAACAGCAGCATTTGATCGTTTGGCATTTTGCATGAGACGAGCTCTCTGTG 1873
QY 61 ThrAlaArgGluAlaValGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTCAGCGGAGGCGAAGAAATTCGCGGAGATGACACAGAACAGCAAGTGATGGA 1933
QY 81 MetLeuGlyAspTrpGluValLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 1934 ATGTGTGGAGATGGAGACATATTAAGCACAGTAGCAAACTCATAGTCGATGTACAG 1993
QY 101 GlyMetProMetAaaIleArgGlyProMetTrpSerValLeuLeuAaaIleGluGluMet 120
Db 1994 GGAATTCCTCATGAACATCCGGGGCCGGGTGTGTACATCTCTCGAATTCAGGAAATC 2053
QY 121 LysLeuLysAaaProGlyArgTyrGlnIleMetLysGluValGlyLysArgSerSerGlu 140
```

```
Db 2054 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAGGCGCAAGAGTCATCGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCAACATGCACTTGAGGTAGAGACGACTCTCCGAAACATATCTTTTAGG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 2174 GATGATATGGAGCGCAAGAGAGAACTATTTCATACATCTCTCGGCTTATTCGGAAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLysSerHisIleAlaLeuPheLeuLysTyr 200
Db 2234 AACCCGAGGTGGGCTCTGACAGGACCTGACCACTACACCGCTTGTCTCTTAT 2293
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCTTAGAGAGAGCAATCTGTGGCACTGTGTGAGCTGTGGCACTGAGAGGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlu 240
Db 2354 CTGCCAGGATTCACACAGCCCAAAATGTGGAGACAGTCCAGGGGCTCCAAAGCCAAAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLysCys 260
Db 2414 CATGTGTACCCCAAGTCAACCAAGACCATGTGGCATGAGGACCAAGAAAGTCTATGC 2473
QY 261 GlnGlnLysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTCTCTGTAGGCTGTCTCTCGGAACCTGATGCGGAACTCTCTGCGG 2553
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 2534 CTCACCTGTGGCGCTGTGGAGGTATTTGTGTGAAGAGAAACAGGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCGCAATGCTCTTAAAGGTTCAGCAGAACCGCTCTGAAAGCATCAGAGTGTGGCTG 2653
QY 321 TrpAlaArgPheCysAaaAaaArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGCAGCTGTGCGGAAACCAATTTCTTGATACCTGGGCCCATGAACATGACACCTGTCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 2714 AAGCATCTTAGGGCTCTTACGAAGAACTTAACAAAGAAAGCAAGGAGCTGCCACCCCA 2773
QY 360 ----- 360
Db 2774 GGCCCAACAGCCCTGGAGCGAAGGTGTGTGACAGAAAGCCCGACGATCTGAACCTGTG 2833
QY 360 ----- 360
Db 2834 GGGCAGTCCCAAGAGCCACCCACCATGCCCAAGCGCTTCCCATGCCAGGACGACAC 2893
QY 360 ----- 360
Db 2894 ACCCTCTCTGTGGATCAAGACACTACAGGCGTGTGTGATGTGACACAGGGGCC 2953
QY 361 ----- 370
Db 2954 ACAACAGAGACCCCAAGAGACTCCAGAGATGACGCAAAACGAGACAGAGGCTCTTGACCC 3013
QY 370 ArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
Db 3014 CAGGCTGTGCGGCTTCAAGTGTGGAAAGCCCTCTGCAAGGGGATATAGGAGGCCCC 3073
QY 390 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPr 410
Db 3074 TCCAGGCGCCACAGCCAGATTCAGAGGCGCCATTGTGTACGTTCCCGGCAAGGGGATC 3133
QY 410 ArgSerSerThrProCysProGlyGlyValAlaValArgGluAspThrTyrProValGlyTh 430
```

```

Db      3133 TCGTTTTCCAGCCCTGTCCTGGTGGGGCTCTCCGGGAAGACAGTACCTGTTGGGCAC 3139
Qy      430 rGInglYvAlProSerProAlaLeuAlaGInglYGlYProGInglYSerTrrPargPhe 450
Db      3194 TCAGGGGTGGCCACGCTCGGCCCCCTGGCTCAGGAGAGACCTCAGGGGTTCCTGAGATTCT 3255
Qy      450 uGInTrPpAnsSerMetProArgLeuProThrAspLeuAspValGluGlyProTrrPheAr 470
Db      3254 GGAGGTGAGAGTCATAATGCCCCGGCTCCCAACGGAACCTGGATTAATAGGGGGGCCCTTGTTCC 3313
Qy      470 gHisrYrAspPheArGInSerCyAsTrpValArgAlaIleSerGInglYAspGInLeuAl 490
Db      3314 CCATTATGATTTTGAACGAGAGCTGCTGGGTCCGTCATATCCAGAGAGACCACTGGC 3373
Qy      490 aProCyArTrpGInAlaGluHisProAlaGlu 500
Db      3374 CACCTGCTGGCAGGCTGAACACTGGCGAGAG 3404

RESULT 4
US-09-620-312D-715
/ Sequence 715, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyun
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yungting
/ APPLICANT: Wang, Duntui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc FL_versions 1.0
/ SEQ ID NO 715
/ LENGTH: 1981
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (59)..(1585)
US-09-620-312D-715

Alignment Scores:
Pred. No.: 4,51e-14 Length: 1981
Score: 344.00 Matches: 150
Percent Similarity: 40.24 Conservative: 77
Best Local Similarity: 26.58 Mismatches: 207
Query Match: 11.58 Indels: 131
Gaps: 26

US-10-071-838-2 (1-549) x US-09-620-312D-715 (1-1981)
Qy      8 GLyserrTPrrP-----AlaGIngluArgGluAspIleIleMetLysrYrGluLys 24
Db      46 GGGCGCGGCTGGAGTGGCGAAGAGCAACGAGAGATAGCGCCGCGCGCGCCGCGCGCG 105

```

QY	253	YHSAAGLAAGLYLeuProGluNarYsgLYProLYsProPhaArg-SerTYrAsnAsp	44
Db	106	GGAAAGCCTGTCGGAAACCCGGGAAGCCTGGCCCAAGGCCCCGAAGCCGCAACCA	165
QY	44	NAhNValAspHisLeuGIYLValHisGIuThrGIuLeuProProLeuThrAlaArgSL	64
Db	166	CGAACTCAGCTCTCTCGGG-----TCGACCTCGAAGGCAACGGCTTCGGCAAGCGCG	219
QY	64	U-----	AlaYgLIuN1 68
Db	220	CATCGACAAGTTCGGCTTATCTGTGGCTCCGAGGGCCCGAGGCGCGCTGGAGAAAT	279
QY	68	eaArgArgLIuLe-----SerArgLYsSerLYrTrValAspMetLeuGLYAspTrpSL	86
Db	280	ACCCCTGAGAGTGTGAGGACAGAGGGAATCCAAAGTGGCTGACATCTCTCAACAATGGGA	339
QY	86	uLYeTYr-----LYsSerSerArgLYsLeuIleAspArgAlaTYrLYsGLYMetProMe	104
Db	340	CAAAAGATGAGCCAAAGAACCAAAAAAATTCCTGTGGCGTGTCCAAAAGGGCATCCCGCC	399
QY	104	CAhN1LeArgGLYProMetCTrPseValLeuLeuAsn1LeGIuGLuMetLYsLeuLYs	124
Db	400	TTCTGTGGGGGCGCTGCTTGGCACTACCTGTGCAGAGGCAAGGTGAATTACAGCAAA	459
QY	124	nProGLYArgTYrGLuIleMetLYsGLuYsGLYAspArgSerSerGIuHis1LeGIuNar	144
Db	460	CCCTGGAAAGTTTGACAGAGCTG--GACATGTCCCTGTGGGACCCCAAGTGTGTGAGCT	516
QY	144	g1LeAspArgAspValSerGLYThrLeuArgLYHis1LePhePheArgAspArgTYrGL	164
Db	517	GATTAGAGGTGACCTGCACCGGACAGTTCCTCATGAATAATGTGTGTCTCCGGGGGG	576
QY	164	YThrLYsGLuNarGLuLeuLYsHis1LeLeuLeuLYaTYrGLuGIuTYrAsnProGLYUva	184
Db	577	CCAGGCGCAAGAGACCTATTCCTGGTGTCTAAAGCCTTACACAGCTGTACCGGCCGAGGA	636
QY	184	1GLYTrCYaArgAspLeuSerHis1LeAlaLeuPheLeuLeuTYrLeuProGLuGL	204
Db	637	GGGCTACACGCAAGGCCGAGGCCCATTCGCGCTGTCTTGGCTCATGACATATGCTCGTGA	696
QY	204	uAspAlaPheTrpAlaLeuValGLuLeuLeuLYaSerGIuNarGHis1SerLeuGLuN1Ph	224
Db	697	GCAGACCTTCTGGTGCTGTGACAGATC-----TGTAAGAAAGTAC--CTGCCCGGCTA	747
QY	224	eHisSerProAsnGLYGLYThrValGLuGLuNarGLuNarGLuN1GLuHis1ValValAla	244
Db	748	CTACAGCGAGAAA-----CTGAGAGCGATCCAGCTGAGCGGAGAGATCCTTTCTC	798
QY	244	aThrSerGLu-----ProLYsThrMetGLYHisGLuNarLYsLYsAspLeuLYsGL	261
Db	799	GCTGTGGAGAAAGGTGTGCGCGGTGGCCCAACAGACCTCAGCCGTCAAGAGATC-----	853
QY	261	YGLuCYsSerProLeuGLYCYsLeuIleArgSLeuIleAspArgLYLeSerLeuGLYLe	281
Db	854	-----GACCGCTCTTATATAGACAGATGTTCATGTGCGCTTCTCCGAACCTT	906
QY	281	U-----ThrLeuArgLeuTrpAspValTYrLeuValGLuGLuGLuGLuN1AlaLe	297
Db	907	GCCTGTGAGCTGTGTGCTGCTGTGGGACATGCTTCTGTGAAGGG-----	955
QY	297	uMetProLeuThrArgLYLea1AspLeuLYValGLuGLuNarGLuNarGLuThrLYsThrSerX	317
Db	956	-----GTCAAGATCATCTTCGCGGTGGGCTG-----	982
QY	317	gCYsGLYProTrpAlaArgPheCYsAsnArgPheValAspThrTrpAlaArgAspGLuAs	337
Db	982	-----	982
QY	337	pThrValLeuLYHis1LeuArgAlaSerMetLYsLYsLeuThrArgLYsGLuGLYAspLe	357
Db	983	-GTCTGTGAAGACAGCGCTGGGCTCCCTGAAAGGTCAAAAGCTGCCAGGGCCAGTA	1041

OY	357	uPro-ProProAlaIalysProGluInglnglySerSerAlaSerArgProValProAlaSerA	377
DB	1042	CGAGACCAATCGAGC-----GACTCCGGAGCCTTCAGCCCCAAGAACA-----	1082
OY	377	rglGlyGlyLysThrLeucCyAlsgslgYlaApArGlnAlaIleProGluIyProProAlaArgP	397
DB	1083	-----TGCAGG-----	1088
OY	397	heProArgProIleTrpSer-----AlaSerPro-ProArgAlaProArg---	411
DB	1089	----AGGCCTTTCTGGTCCAGGAGGTGTGTGAATTGCCCCGAGAACAGACCAGATTGA	1143
OY	412	SerSerThrProCyAProGlyGlYAlaValArgGluAspThrTyPrProValGlyThrsIn	431
DB	1144	GCGCAAAACCTCATTCAGCTTGCGGGCTGCGCAGGAGACCCG-----GGGTGAGCT	1194
OY	432	GlyValProSerProAlaIalaIaglnInglyGlyProGlnlglySerTrpArgPheLeuGln	451
DB	1195	GCAGTGCCTCCGCCGCCAGGCT--GCANGTCCCAAAGGTATCTTGATGCAGAAAC	1251
OY	452	TPPaSermEProArgLeuProThrArAspleuAspValGluGlyIyProTrpPheArgHts	471
DB	1252	TGRTCC-----CCGGCCCTGCCCTTCAACCTTCACACATTCATCCGCTGCCCTAGA---	1302
OY	472	TyrAspPheArgInsErCyS-----TrpValArgAlaIleSerGlnGluAspGln	488
DB	1303	-----TGCCCCCTCCCTGGCTCCAAAGCCCAAGCCCAAGCCCAACCA	1344
OY	489	LeuAlaProCyTrpGlnAlaIaglnHstProAlaGluArgSerFlaPheAlaIa	508
DB	1345	GCAGGCCCAAGAGCACGCGAA--ACAGATGAAGGGAGAGGGCAGCTGGAGAAAC	1401
OY	509	ProSerThrApSerAspGlnInglyThrProPheArgAlaArgAspGluGlnPro-----	526
DB	1402	CCCAGCCCAATCAAGCATGTGTGTGGTGCCTGCAGAGATGATGTCCCCCAACA	1461
OY	527	CyAlaIaPro 529	
DB	1462	TGTGCCCC 1470	
RESULT 5			
US-09-774-528-336			
; Sequence 336, Application US/09774528			
; Patent No. 6743619			
GENERAL INFORMATION:			
APPLICANT: Tang, Y. Tom			
APPLICANT: Zhou, Ping			
APPLICANT: Goodrich, Ryle			
APPLICANT: Liu, Chenghua			
APPLICANT: Asundi, Vinod			
APPLICANT: Ren, Feiyang			
APPLICANT: Zhang, Jie			
APPLICANT: Zhao, Qing A.			
APPLICANT: Yang, Yonghong			
APPLICANT: Xue, Aidong J.			
APPLICANT: Wehrman, Tom			
APPLICANT: Wang, Jian-Rui			
APPLICANT: Wang, Dunrui			
APPLICANT: Drennack, Radjoje T.			
TITLE OF INVENTION: No. 6743619el Nucleic Acids and			
FILE REFERENCE: Polypeptides			
CURRENT APPLICATION NUMBER: US/09/774, 528			
CURRENT FILING DATE: 2001-01-30			
NUMBER OF SEQ ID NOS: 441			
SOFTWARE: pc_FL_genes Version 2.0			
SEQ ID NO 336			
LENGTH: 1823			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (136)..(1476)			

```

Db      993 -----TGTGG-----998
Qy      326 AsnArgPheValAspThrTrpAlaArgAspGlnAspThrValLeuIleuValAla 345
Db      999 -----GCTGACACTGGTGGCGCTGGCGCTGGG-----CACTGGACAGACA 1037
Qy      346 SerMetIleValSerThrArgIleValGlnGlyAspLeuPro---ProProAlaIleuProGlu 364
Db      1038 GCCAGG-----GGCTGCTGGCGCTGGCTGGCA-----GAC 1067
Qy      365 GlnGlySerSerAlaSerArgProValProAlaSerArgIleGlyIleThrLeuGlyLeu 384
Db      1068 ACTGGGAGCGCTTCGAGCCATCCCGCGGACAGCTGCAAGAGAGGCGCTTCATGTCACA 1127
Qy      385 GlyAspArg-----GlnAlaProProGly 392
Db      1128 GGTGCACAGCGTGCTGTCTGACAGCGGACCTGACGCGGAGATCAAGGCCCGAGCTGGC 1187
Qy      393 ProProAlaArgPhe-----ProArgProIleTrpSerAlaSerProProArg 408
Db      1188 CCAGCTGCCGATTCGCGCGCGGACCCCGCCCG-----GCCACAGGTCCGCTCGC 1241
Qy      409 -----AlaProArgSerSerThr-----414
Db      1242 CGGGGCCCAAGCCATCTTGAAGCCCGACAGCTGGACAGATGCGACAGAGCCCAAGCC 1301
Qy      415 -----ProCysProGlyValAla-----420
Db      1302 TGAAGTGCTCGGATTGTGTGTGACAGCCCGGAGAGGCCAGACCAAGCGGCGGAAACC 1361
Qy      421 -----ValArgIleAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeu 438
Db      1362 CCAGACCCCGGCGAAGAC--TTTCCATGGGCTCTGACTCGGAGCCCGGAGCCCGCC-- 1415
Qy      439 AlaGlnGlyValProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeu 458
Db      1416 CATCGAGGGGCGCCCGCAGGCCCAAGAG-----CTCCACCTCTT 1457
Qy      459 ProThrAspLeuAspValGlnGlyProTrp-----468
Db      1458 CCTGGACACCCGCTTGAAGAGACATGAGACTTAAGTCCCGCAGTCAATTCCTGA 1517
Qy      469 -----PheArgHisTrpAspPheArgGlnSerCysTrp 479
Db      1518 TGGCTGATGCCAGCCCGGCAATAGGACCGCACTTACTCTTGGACCTCGGGAATTGG 1577
Qy      480 ValArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTrpGlnAlaGlnIleProAla 499
Db      1578 CTT-----CCTTCCTGGCAAGGACCAAGCAATGGG 1607
Qy      500 GlnArgValArgSerAlaPheAlaAlaProSerThrAspSerAsp-----514
Db      1608 GAAAGGAGG-AGGTCTCCGCTGTAACATAGTGGTCAAGGACCTAGCATGAGAGGGTCAAC 1666
Qy      515 -----GlnGlyThrProPheArgAlaArgAspGlnGlnProCys 527
Db      1667 AAGATGGGACCTGAGAGACCCATGGAACCGTCCGTGGCCAGGCGCTTCACAA----- 1720
Qy      528 AlaProThrSerGlyProCysIleuGlyLeuIleGlnIleSerGlnPheProPro 547
Db      1721 GTACCAAGCCAGACCAACAAAGAGTCAAGGAGGGTTGGCTGATGTCAGAGGCCCAAGA 1780
Qy      548 Gly 548
Db      1781 GGG 1783

```

RESULT 6
 US-10-120-988-336
 ; Sequence 336, Application US/10120988
 ; Patent No. 6919193
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom

```

; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Felyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Zemanac, Radolje T.
; TITLE OF INVENTION: No. 6919193el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pc FL genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
; US-10-120-988-336

```

```

Alignment Scores:
Pred. No.: 1,2e-13 length: 1823
Score: 337.00 Matches: 157
Percent Similarity: 36.2% Conservative: 68
Best Local Similarity: 25.3% Mismatches: 231
Query Match: 11.3% Indels: 167
DB: Gaps: 29

```

US-10-071-838-2 (1-549) x US-10-120-988-336 (1-1823)

```

Qy      16 GlnAspIleIleMetIleValGlu---LysGlnIleArgAlaGlyLeuProGluAsp--- 33
Db      154 GAGAGCTGGTGGTGAAGCTCCCGAGCTGACAGATGATCCAGCTCTGGGGTCCGACTCA 213
Qy      34 -----LysGlyProIleProPheArgSerTrpAsnAsnValAspHisLeuGlyIle 51
Db      214 GAGCTCAGGGGCGCTGGCCCATATCC-----CAGGCCACCGCTATGATTC 261
Qy      52 ValIleGluThrGluLeuProPheLeuThrAlaArgGluAlaValGlnIleArgGlu 71
Db      262 ATTGGGGGCACTCAGACAGACCCAGGCGGCGGACCCACCTCAGACCTCATCCGCCAA 321
Qy      72 IleSerArgIleSerIleTrpValAspMetLeuGlyAspTrpGluIleValSerSer 91
Db      322 -----CGGAGATGAAGTGGGTGAGATGACTCGACCTGGAGAAACCATGTCGG 375
Qy      92 Arg-----LysLeuIleAspArgAlaTrpLysGlyMetProMetAsnIleArgGlyPro 109
Db      376 CGGTACAGAAAGGTAAAGATGACAGTCCCGGAAAGGACATCCCTGCTCCGCGCCGA 435
Qy      110 MetTrpSerValIleLeuLeuAsnIleGluGluMetLysLeuIleValAsnProGluArgTrpGln 129
Db      436 TGCTGGCCCTGTGTGTGTGGGCGCCATGTGTGCAAGAAAGCCCTGGACCTATCAG 495
Qy      130 IleMetLysGluIleValArgSerSerGlnIleIleGlnArgIleAspArgVal 149
Db      496 GAGCTGCAAGG---GCCCTGGAGACCCAGATGATGAGACCATTTGGACAGGACCTG 552
Qy      150 SerGlyThrLeuArgIleValIlePhePheArgValTrpGlyThrLysGlnArgGlu 169
Db      553 CACCTCAATTCCTCTGACAGATGTTGTGCTCCCTCAGGCGGACCGGCGGAG 612
Qy      170 LeuLeuIleIleLeuLeuAlaTrpGluGluTrpAsnProGluValAlaGlyTrpCysValAsp 189
Db      613 CTCCTCAGAGTGTCAAGGCTTACACCTGTATGACCCGAGAGAGGCTTATCCAGGCC 672
Qy      190 LeuSerHisIleAlaIleLeuPheLeuLeuTrpGluProGluGlnAspAlaPheTrpAla 209
Db      673 CAGGGGCGGCTGTGTGTGTGCTGCTCATGACCTGCGCCGAGAGAGGCGCTTCTGTGTC 732

```

QY	210	leuValGlnLeuLeuAlaSerGlnThrArgHisSerLeuGlnGlyPheHisSerProAlaGly	229
Db	733	CTGGTGCAGATGCT-----GAGGTACTCTCCCTGGTACTACGGGGCCAC---	780
QY	230	GlyThrValGlnGlyLeuGlnAAserGlnGlnGlnHisValAlaAlaThrSerGlnProLys	249
Db	781	-----ATGGAGGCTGTGGCGGCTGGAGCGCCGAGAGGTTCATGAGCGCTGGCGG---CGG	831
QY	250	ThrMetGlyHisGlnAAsrLysLysAAsrLysCysGlyGlnCysSerProLeuGlyCysLeu	269
Db	832	CTGCTTCCGCACTGGCAGACAGACCTGTCAGACAGTGGAGGTGGAGACCCCTGCTGTAACCTG	891
QY	270	-----IleArgIleLeuIleAAsrGlyIleSerLeuGlyLeuThrLeuArgLeu	285
Db	892	CCCGAGGTGCTCTGGCTCTTGGCTCTTGGCGGCTCTGCTCCCTTCCCAAGATGCTGGTGGTC	951
QY	286	TrpAAsrValTyrLeuValGlnGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPhe	305
Db	952	TGGGATGCTTCTCTCAGTGAAGGTGC-CAGAGTACTGTTCCG-----	992
QY	306	LysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaAAsrPheCys	325
Db	993	-----TGTGGG-----	998
QY	326	AAsnArgPheValAAsrThrTrpAlaAAsrArgGlnAAsrThrValLeuLysHisLeuArgAla	345
Db	999	-----GCTGACACTGTGTGGCTGGCTGGCTGGG-----CAGTCGAGAGCA	1033
QY	346	SerMetLysLysLeuThrArgLysGlnGlyAAsrLeuPro---ProProAlaLysProGlu	364
Db	1038	GCGAGG-----GGCTGCTCTGGCTCTCTGGA-----GAC	1066
QY	365	GlnGlySerSerAlaSerArgProValProAlaAAsrArgGlyLysThrLeuCysLys	384
Db	1068	ACTGGAGAGCCCTTCAGAGCATCCCCCGCAGAGTCAGAGAGAGAGCCCTCATGTACACA	1127
QY	385	GlyAAsrArg-----GlnAlaAAsrProGly	392
Db	1128	GGTGCAGACCGTGTGCTGTGCAGAGCGGAGCACTGCAGCGGAGATCAAGGCCCACTGAGC	1187
QY	393	ProProAlaArgPhe-----ProArgProIleTrpSerAlaSerProArg	408
Db	1188	CCAGCTGCCCATTTCCGGCCGGGACCCCGCCCG-----GCCACAGTCCGCTGCC	1243
QY	409	-----AlaProArgSerSerThr-----	414
Db	1242	CGGGGCCCAAGCCATCTTGAAGCCCAAGACGCTGGAGAGGTCGACGAGGGCCCAAGCC	1303
QY	415	-----ProCysProGlyGlyAla-----	420
Db	1302	TGAGGTGCTCGAATTGTGGTGCAGCCCGCCGAGAGGCCACAGCACCGCGCGGAAAC	1361
QY	421	-----ValArgGlnAAsrThrTyrProValGlyThrGlnGlyValProSerProAlaLeu	438
Db	1362	CCAGACCCGGGCAAGAC---TTTCCATGGGCTCTGACTCGGGCCCGGGGCCCCC---	1415
QY	439	AlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAAsrMetProArgLeu	458
Db	1416	CATCAGAGGGGCCCCCAGAGGCCCAAGAGG-----CTCACCTCTT	1455
QY	459	ProThrAAsrLeuAAsrValGlnGlyProTrp-----	468
Db	1458	CCCTGACACCCGCTTCTGAGAGACCAATGAGCTTAAGTGTCCCCAGCTCATATTGCTGGA	1517
QY	469	-----PheArgHisTyrAAsrPheArgGlnSerCysTrp	479
Db	1518	TGGCTGATGCCAGCCCGGCAAAATAGGACCCGACCTTTACTCTTGGGACTCGGGGACTTGG	1577
QY	480	ValArgAlaIleSerGlnGlnAAsrGlnLeuAlaProCysTrpGlnAlaGlnHisProAla	499
Db	1578	CTT-----CCTTCTCGGCAAGACACGAGAGTGG	1607
QY	500	GlnArgValArgSerAlaPheAlaAlaProSerThrAAsrSerAAsp-----	514

```

Db      1608 GAAAGAGC-AGGTCCTCCGTGGTACTACTGGGTCAAGGCACATACGAATGAGAAGGGTCAAC 1666
QY      515 -----GlnGlyThrProPheArgAlaArgAspGluInProCyS 527
Db      1667 AGAGTGCGGACAGTGAGAGCACCATGGAACCGTCGTGGTCCCAGAGGCCCTCACAAA----- 1720
QY      528 AlaProThrsSerGlyProCyLeucyAspGlyLeuHisIleuGluSerGlnPheProPro 547
Db      1721 GTACCAGAACCCAGACACCAAAGGATCGAGGGAAGGGGGTTGGCTGAGTCAAGGGACCCCAAGA 1780
QY      548 Gly 548
Db      1781 GGG 1783

RESULT 7
US-09-949-016-2139
; Sequence 2139, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2139
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2139

Alignment Scores:
Pred. No.:          2,596-13           Length:          1925
Score:              332,50             Matches:          119
Percent Similarity: 39.6%               Conservative:     57
Best Local Similarity: 26.8%            Mismatches:      135
Query Match:        11.1%               Indels:          133
DB:                  3                 Gaps:           15

US-10-071-838-2 (1-549) x US-09-949-016-2139 (1-1925)
QY      113 ValIeuAeuAmIleGlunGluMetLysLeuYsaenProGlyArgTyrr--GlnIleMet 131
Db      30 GTAGACATTCAAACTTGCCACGAGGAGGAGCTGAATAATGTTGTAATAATATGACAGAAATTA 89
QY      132 LysGluTybGlyLysArgSerSerGlnHisIleGlnArgIleAspArgSerValSerGly 151
Db      90 AAACAACAGACACGGGGCTGTTCACCTGCATGCACAGACAAATAGACCTGATGTCAACCCG 149
QY      152 ThrIeuArgLysHisIlePhePheArgAspArgTyrgLyThrLysGlnArgIleuLeu 171
Db      150 ACATTTCGGGACACCATATATGTTTAGAGACAGATATGGCTTTAAGCAACAATCCTTATTC 209
QY      172 HisIleLeuLeuAlaTyrgLunGluTyraenProGluValGlyTyrcyAspArgAspleuSer 191
Db      210 CATGGCTTGCTGGCTTATTCATTATTAACAAGGAAGTCGGGTATATGTCAGGGGATGAGC 269
QY      192 HisIleAlaAlaLeuPheLeuLeuTyrlenuProGluGluAspAlaPheTrpAlaIleuVal 211
Db      270 CAGATCAACAGCTTTATCTCCTCATGTATATGAACGAGGAAGAGCCCTTCGGGCCCTGGTC 329
QY      212 GlnIleuAlaIasrclJnArgHisSerIleuGlnGlyPheHisSerProAsnGlyGlyThr 231
Db      330 AATCTTTCTCAGGCCCTTAACATCAGATG----- 359

```

```

QY      232 ValGlnGlyLeuGlnAserGlnGlnValAlaThrSerGlnProLeuThrMet 251
Db      360 -----CATGAACATCTAATGAATTGCTCATGAAGA----- 392
QY      252 GlyHisGlnAserLeuValAserLeuGlyGlnCysSerProLeuGlyCysLeuIleArg 271
Db      392 ----- 392
QY      272 IleLeuIleAserGlyIleSerLeuGlyLeuThrLeuArgLeuThrAserValTyrLeuVal 291
Db      393 -----CTTGTA 398
QY      292 GluGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLeuValGlnGlnIleArg 311
Db      399 GAA----- 401
QY      312 LeuThrLeuThrSerArgCysGlyProTrpAlaArgPheCysAserArgPheValAserThr 331
Db      402 -----TTTTTCAGAGACC 416
QY      332 TrpAlaArgAser-----GluAserThrValIleuValHisLeuArgAlaSerMet 347
Db      417 CTGGCAAAAGATTTTCTTTGAAGATGATTTGCTGATAGCAACTTCAGATTTCTATG 476
QY      348 LysLeuLeuThrArgLysGlnGlyAserLeuProProAlaLys-----Pro 363
Db      477 ACGAACTAAAGCGGCAAGTTAGACTTCCAGAACCTGTAAAGAGATGAATATCCA 536
QY      364 GluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCys 383
Db      537 AAGAAGCCCTGGGGGCGAGCTTCCAGCTTCAAGCTTCGTTGGGCGCTCCATCACTTGAGC 596
QY      384 LysGlyAserArgGlnAlaProProGlyProProAlaArgPhePro----- 398
Db      597 AACGGACAGAGAGCCGT-----GGCCGGCCGAGCCGCTGGCCAGCCGAGAGAGGAG 650
QY      399 -----ArgProIleTrpSerAlaSerProProAlaAlaProArgSerThrProCys 416
Db      651 AGCGGGGCGCCCAACAGAGAGCAGAGCACTCCCGCAGCCCAAGAGAGAGAGAGAGAGC 710
QY      416 sProGlyGlyAlaValArgGlnAser-----ThrTyrProValGlyThrGlnGlyValPc 434
Db      711 CCCGAGAGAGCAGAGCCGCCAAGACCGAATCGGTGAGAGAGAGAGAGAGAGAGAGAGAG 770
QY      434 oSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAse 454
Db      771 GATGAGGACAGATTTTCAAGAAACTCCATCGGCTCCACAGAGACAGTTCAGAGCAATAT 830
QY      454 rMetProAserLeuProThrAserLeuAserValGlnGlyProTrpPheArgHisTyrAser 474
Db      831 AATCAGCAGCTGCTCAACAAATATAGCA-----CGCCACTTCMAAT 872
QY      474 heArgGlnSer-----Cys-----TrpValArgAlaIleSerGlnIleAserG 488
Db      873 ATCAGAGAGAGAGTTTGTGCCCCAATGAATGAACCGTCAGACGCTTCAGTACAGAGAGA 932
QY      488 IleuValaProCysTrpGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAla 508
Db      933 AC-----TGC-----CAAAATACACATGAGAGCAAGTGCAGAGAGAGAGAGAGAG 980
QY      508 Ia-----ProSerThrAserAserArgGlnGlyThrProPheArgAlaArgAser 524
Db      981 CTCGCAAGTTACCGTCCAGGTCTGCGAGAGTGGGAGTGTCAAACTGCGGCAAGAGATG 1040
QY      524 IuGlnPro 526
Db      1041 AAGGCCCT 1048

```

```

; GENERAL INFORMATION:
; APPLICANT: Tang, X. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aildong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhimei
; APPLICANT: John Tillinghaet
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_Fl_genes Version 1.0
; SEQ ID NO 675_
; LENGTH: 3039
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (169)..(2502)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3039)
; OTHER INFORMATION: n = a,t,c or g
; US-09-620-312D-675

Alignment Scores:
Pred. No.: 4,97e-08 Length: 3039
Score: 258.00 Matches: 157
Percent Similarity: 35.9% Conservative: 77
Best Local Similarity: 24.1% Mismatches: 236
Query Match: 8.6% Indels: 183
DB: 3 Gaps: 30

US-10-071-838-2 (1-549) x US-09-620-312D-675 (1-3039)
QY      9 SerTrpTrpAlaGlnGlnIuArgGlnAserIleIleMetLysTyr---GluTyrGlyHisArg 27
Db      220 AGCATATGCCCCCAGAG-----ATCTTGCCAAATACACGAGAGAGAGAGAGTCA 270
QY      28 AlaGlyLeuProGln----- 35
Db      271 GCAGAGCAACAGAGTTCTACTACATGATGATGTTGTTCCGCTGTGTACAAAGAGAAAGGT 330
QY      36 ProLysProPheArgSerTyrAserAserValAser----- 47
Db      331 GATGAGCTGGCTGCTCCAGTGTGCTGGCACTCCCTCTGATGAGAGATCTCCAGAGG 390
QY      48 -----HisLeuGlyIleValHisGluThrGluLeuProProLeuThrAla 62
Db      391 CTCGGGTGAGAGCCCACTGAGATTCAACCATTAACAGATGTGGGATTCACCC----- 447
QY      63 ArgGlnAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAserPheLeu 82
Db      447 -----TyrLysSerSerArgLysLeuIleAserArgAla 98

```

RESULT 8
 US-09-620-312D-675
 ; Sequence 675, Application US/09620312D
 ; Patent No. 6569662


```

/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (169)..(2418)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(2955)
/ OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-676

```

```

Alignment Scores:
Pred. No.: 1,04e-07 Length: 2955
Score: 253.00 Matches: 156
Percent Similarity: 35.8% Conservative: 71
Best Local Similarity: 24.6% Mismatches: 233
Query Match: 8.5% Indels: 175
Gaps: 28

```

US-10-071-838-2 (1-549) x US-09-620-312D-676 (1-2955)

```

QY 9 SerTPTRAlaGlnGluArgGluAspIleIleMetLysTyr--GluYsgIyHsArg 27
DB 220 AGCATATGCGCCCGAGAG--ATCTTGCCAAAGTACAGAGAAAGAGTCA 270
QY 28 AlaGlyLeuProGlu-----AapYsgIy 35
DB 271 GCAAGACAAACAGAGTTCTACTACAGATGATTGGTTCCGTGTGTAACAAGAAAGGT 330
QY 36 ProLYsProPheArgSerTyrAsnAsnValAsp----- 47
DB 331 GATGAGCCTGGCTCCAGTCTGCGGGACTCCCTCTGATGAGAGATGCTCCAGAGG 390
QY 48 -----HisGluYIleValHisGluThrGluLeuProIleThrAla 62
DB 391 CTGGCGTGACAGGCCACCTGAGATTCAACCATACACAGATGTGGGGATCTCACC--- 447
QY 63 ArgGluAlaLeuGlnIleArgArgGluIleSerArgLysSerIystrValAspMetLeu 82
DB 447 ----- 447
QY 83 GlyAspTRGluLys-----TyrLysSerSerArgLysLeuIleAspArgAla 98
DB 448 -----TGGACAAAGATTGCCGTCTCCCTACCCCGCTCGAAGAGCTCCGCTCCGCGTG 501
QY 99 TyrLYsGlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGlu 118
DB 502 CTGGCCGGCATCCCATGTCATGAGGCCACAGCTGTGATGCGGCTCTCGGGGCCCTG 561
QY 119 GluMetLYsLeuLysAsnProGlyArgTyr---GlnIleMetLysGluYsgIyLysArg 137
DB 562 CAGAAAGAAAGAACTCTAGCTGTCTACCGGAAATTGTGAAGAACAGCTTCAACGAT 621
QY 138 SerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIle 157
DB 622 GAGACCATGCTGCCAAGACAGATCGAAGAGACTCTCCGCAACCTAGCCCAAGACGCC 681
QY 158 PhePheArgAspArgTyrGlyThrLYsGlnArgGluLeuLeuHisIleLeuLeuAlaTyr 177
DB 682 TGCCTTCGACAGATGGGTAGCATCGGGGTGCCCGCTCGCAAGGTGCTCGGGGCCCTG 741
QY 178 GluGluTyrAsnProGluValGlyTyrCyAspArgLeuSerHisIleAlaIleLeuPhe 197
DB 742 GCCTGGCTTACCCCAAGATCGGTACTCTCCAGGGACACCGCATGGGTGGCGCTGCCCTC 801
QY 198 LeuLeuTyrLeuProGluGluAspAlaPheTyr-----AlaLeuValGln---Leu 213
DB 802 CTGCTGTCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
QY 214 LeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGln 233
DB 862 CTCGCCGCTCTACTTACAGAC---ACCTGCGCTG 894

```

```

QY 234 GlyLeuGlnAspGlnGln-----GluHisValAlaIleHisSerGlnProLYsThr 250
DB 895 GGTGTCCAACTACACAGAGGGGTCTCCGCCACCTATTGTCCAGTAACCTGCTCCGCTG 954
QY 251 MetGlyHisGlnAspLYsAspLeuCYsgIyGlnCYsSerProLeuGlyCYsLeuIle 270
DB 955 GACAACTGCTCCAGAGCATGATCATTTAGAGCTGTCCCTGATCATCACTGCAGCTGTTCTC 1014
QY 271 ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuThrAspValTyrLeu 290
DB 1015 ACGGCTTGCCAGCGCTGGTGGATCAATCAAGCTCTCTGCGCATCTGGGACCTGTTTTC 1074
QY 291 ValGluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLYs 310
DB 1075 TACGAGAGGTCCCGGTGCTTTCACCTCACTGAGCTGGGATGCTGCACCTCAAGAGGAA 1134
QY 311 ArgLeuThrLYsThrSerArgCYsgIyProTRPAlaArgPheCYsAsnArgPheValAsp 330
DB 1135 GAGCTGATCCAGTACAGAAACTCG-----GCTTCATCTTCAACACGCTATGGAT 1185
QY 331 ThrTRPAlaArgAspGluAspThr-----ValLeuLysHisIleAspArgAla 345
DB 1186 ATCCCGCTCCAGATGGAGGACCGGAGCTGTTCTGGGGGTGGCCATGCGGCTGGCGGC 1245
QY 346 SerMetLYsLeuThrArgLYsGlnGlyAspLeuProProAlaLysProGluGln 365
DB 1246 TCCCTCAC-CGATGTGGCT---GGAAGCTACGCCCGGAACT---CCT 1289
QY 366 GlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCYsLYsGly 385
DB 1290 GGCCTATCTCATTTGAGAGACAGAGGCGAGT----- 1319
QY 386 AspArgGlnAlaProProGlyProProAlaArgPheProArgProIleTRPserAlaSer 405
DB 1320 -----CCTGGGGGCCGACCCCTCACCAACCTCTCAGGTTGT 1358
QY 406 ProProArgAlaProArgSerThrProCYsProGlyValAlaValArgGluAspThr 425
DB 1359 TCGCCGAGAGACCCAGCGAGAGATCCACATCACTGCTGTGCTTTCG---GGAGCA 1415
QY 426 TyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyLYs---ProGln 444
DB 1416 TGACCTGAGAGGACCTCAAGAGCAAGAAATCAAGAGAGAGAGAACTGTGTGACCTCGG 1475
QY 445 GlySerTRP-----ArgPheLeuGlnTRP 452
DB 1476 GGAAGCATCTCGCGGTGGACGCGCATTTCCAGTGCAAGACCCCAAAACTGACAGCGT 1535
QY 453 AsnSerMetProArgLeu-----ProThr-Asp----- 461
DB 1536 GAGCTGACTCCAGACTATAGCATGAGAGCCACGAGCGGACACGAGAACTACGTGTC 1595
QY 462 -----LeuAspValGluGlyProTRPPh 469
DB 1596 GTGCTACGACACCAACCGCGCGGACCAAGCCCTGCTGAGCTTTGAG----- 1644
QY 469 eArgHisTyrAsp-----PheArgGlnSerCYstrValArgAlaIleSerGly 485
DB 1645 -CGGACGAGACGACGAGAGCTGGCTCCGCAAGAACATCATCATCAATGCTGCTCA 1703
QY 485 nGluAspGlnLeuAlaProCYstrProGlnAlaGlu-----HisPhe 498
DB 1704 GAAAGACAG---CACTGCTGGGTGGGAGGTCAACGAGCTGCGAGGCTGTTTC 1757
QY 498 AlaGluArgVal-ArgSer---AlaPheAlaAlaProSerThrAspSerArgGlnLYs 517
DB 1758 AGCCAAAGTTCTGGAAAGTCTTGATGAGAGGACAAAGAGTCTCATCTCGCGGGAGTCA 1817
QY 517 hPProPheArg-----AlaArgAspGluGlnP 526
DB 1818 CTCGGTGAACGAGGGGTACACAGACTGTCGAGAGGACCCCTGCTCCGCGCTTAAGGC 1877
QY 526 roCYAlaPro-----ThiSerGlyProCYsLeu 535

```



```

QY 95 ILeapArgLaTyrlYsGlyMetProMetAsnIleArgGlyProMetTrpSerValIleu 114
DB 492 AAGGAACCTTGCTTAAGGATACCCCATCTTAGAGCAATAGTTGGCACTTTA 551
QY 115 LeuAenIleGluGlyMetLysLeuLysAsnProGlyArgTyrlGlnIleMetLysGlyLys 134
DB 552 TGCATGTCACAAAGTATGCCAATTAAAGATCAG-----TATTCAGAACATC 596
QY 135 GlyLysAsnSerSerGlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArg 154
DB 597 CTGAAAATGACCTCGCTTGAAAAAATGATCCGAAAGGACATTCGTAACCTTAACCT 656
QY 155 LysHisIlePhePheArgAspArgTyrlGlyThrLysGlnArgGlyLeuLeuHisIleLeu 174
DB 657 GAACACAACTTTTAAAGAAAAAGTACCTTGACAGAGAGTTTATTATTAATGAATG 716
QY 175 LeuAlaTyrlGluGlyLysAsnProGlyValGlyTyrlCysArgAspLeuSerHisIleAla 194
DB 717 AAGGCTTACTCTTATGATGATCGTGAAGTGGTACTGTCAGAAAGTCTTTTATATG 776
QY 195 AlaLeuPheLeuLeuTyrlLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeu 214
DB 777 GGATGTGCTTATGAGATGCCAGAAAGAAAGACTTCTGTGTATTTGTTAAATTAATG 836
QY 215 AlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly 234
DB 837 CAAGATTATGA-----CTTCGTGAACCTTTTAAACCAAGTATGGCAGAAATGGCCCTT 890
QY 235 LeuGlnAspGlnGlnGluHisValAlaAlaThrSerGlnProLysTrpMetGlyHisGln 254
DB 891 TGTATGATACAGTGTGAATGATGATACAGAGACATCTTCAGAGCTCTTTCAGACTTTT 950
QY 255 AspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIle 274
DB 951 CAATCTCAGAGTTTTCATACCTCAATGATGATGATCATCTCGTTTCTGATCTTCTTCT 1010
QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrlLeuValGlnGlyLeu 294
DB 1011 ACACTTTTCCACTACCACTGTCACACAGATATTTGATATCTTATATGCTGAGGCTTTA 1070
QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlyLysArgLeuThrLys 314
DB 1071 GAAATAGTGTTCGTGATGAGATTAACAATCTTCAGATGATCAGGACAGACTGATGCA 1130
QY 315 ThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArg 334
DB 1131 CTTCAGATGAAAGGATGTTACAGACATTTCAAAAGTCATTCACATCAGTTGATGCT 1190
QY 335 AspGluAspThrValLeu-----LysHisLeuArgAlaSerMetLys 348
DB 1191 GTCCAGACAGACTTATCCAGACAGCTTACCAAGTCAATACATTCAAAAAAATGAAA 1250
QY 349 LysLeuThrArgLysGlnGlyAspLeuProProAlaLysProGluGln 365
DB 1251 AAGCTTGAAGAAAGATATACCTACATTAATAAAGAAAGAAATGAAAGAGCA 1301

```

```

Alignment Scores:
Pred. No.: 7,13e-06 Length: 1887
Score: 222.00 Matches: 129
Percent Similarity: 36.6% Conservative: 64
Best Local Similarity: 24.5% Mismatches: 198
Query Match: 7.4% Indels: 138
DB: 3 Gaps: 23

US-10-071-838-2 (1-549) x US-09-270-767-12482 (1-1887)
QY 47 AspHisLeuGlyIleValHisGlyLysThrGlyLeuProProLeuThrAlaArgIleLysLys 66
DB 105 GACCTTTCGCTCGAGTTTCAT-----GTTCTTTAGTCGCGAAGCCGCTAAG 152
QY 67 -----GlnIleArg 69
DB 153 TACGACATTCCTCGAGACAGACAGCCGCTGATGAACACATTCAGACACAGATTACAGC 212
QY 70 ArgGluIleSerArgLysSerLysTrpValAspMetLeuGlyAspTrpGluLysTrp--- 88
DB 213 GCAGAGATCATACAGAACAGAGAG-----GAAAAGATGTGCGCTGGAGGCGGCAATTTC 266
QY 89 -----LysSerSerArgLysLeuIleAspArgAlaTyrl 99
DB 267 CCGGACTTCGAGCGGGGCAATTGGAATGTTCCGACACAGAGACTGATTAATCATATAGT 326
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuAsnIleGluGln 119
DB 327 GAGGCTATTCCTGACCAAACTGCGCAGAGATCTGCTGATATTTCTCGGGGGCAATCCAC 386
QY 120 MetLysLeuLysAsnProGlyArgTyrlGlnIleMetLysGlyLysGlyLys-----Lys 136
DB 387 GACAAAGAGATGAATCCGAGCTGTACAGAGATCTCGTGAAGAAAGCGGCTTGCAATAAG 446
QY 137 ArgSerSerGlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHis 156
DB 447 AACTGCTTGGCCAC---GATGAATATGATCGGACTTGGCCGCTTCGCTGCCAGACAT 503
QY 157 IlePhePheArg-----AspArgTyrlGlyThrLysGlnArgGlyLeuLeuHisIleLeu 174
DB 504 CCGGATTTTCAAAAGACAGCGATGGCATAGTGTCTGCGAGAG-----GTCTT 551
QY 175 LeuAlaTyrlGluGlyLysAsnProGlyValGlyTyrlCysArgAspLeuSerHisIleAla 194
DB 552 CAGGCTTATGCGCTGCGGAATCCACAGATGGGCTACTGTCAAGCCCATGAATATGTGTG 611
QY 195 AlaLeuPheLeuLeuTyrlLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeu 214
DB 612 TCGGTGTTCTGTCTTCTGTGATGAGAAAGCATTTCTGAGTCTGCGCAGCCTGTGTC 671
QY 215 AlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly 234
DB 672 -----GAAACTTACTACCGGACTTACCTACCAAGATATA-----GTGTTGGGC 713
QY 235 LeuGlnAspGlnGln-----GlnHisValValAlaThrSerGlnProLysTrpMet 251
DB 714 GCCCAATGATCAGGGGTGCTTAATGAGCTAGAGAGAGCATTTGCTGATTTGCAAC 773
QY 252 GlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArg 271
DB 774 GGCCATCTCGAGCACTTGGGTGTATTAAGATGATTTCC---ATTTCCTGTCTTCTTAC 830
QY 272 IleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrlLeuVal 291
DB 831 ATCTTCATGACGCGATCAGCTACGAGAGCTGCTGACATCTGCACTGCTTTTATTC 890
QY 292 GlnGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLysArg 311
DB 891 GAGGCGCCCAAGATCATTTTATGATGATTCATGCAATCATGATGATGAAAGAGAGCAAA 950
QY 312 LeuThrLysThrSerArgCysGlyProTrpAlaArg-PheCysAsnArgPheValAspThr 331

```

```

RESULT 12
US-09-270-767-12482
; Sequence 12482, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OR INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12482
; LENGTH: 1887
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12482

```

```
Db 951 CTGTTGATCTGCCAAGATGATGCGGAAGCAATGCTAGTTTTCAGA-----ACTAC 1001
Qy 331 rTPRAlaArgAspGluAspThrValLeu----- 340
Db 1002 CTGAGG-----GAGTTTACAAATCCAGACTACAGGTGCTCTCTACGACGAAACGG 1055
Qy 341 -----LysH1sLeuAlaArgLaserMetLysLeuThrAr 352
Db 1056 AAGATGAGGGTAAAGTCCAAACAAACAGTTACAGCCGTGATTCATGAGGATACACA 1115
Qy 352 gLygInGlyAspLeuProProAlaLysProGluInGlySerSerAlaserArp 372
Db 1116 AAAAGATTCCGAT-----GAGGAGATATCACACAGCAGCGCATCGAGAG 1160
Qy 372 ovalProAlaSerArgLysGly----- 379
Db 1161 TTAAGCAACAGACATCGCCGGTTAAAGATGCGACAGATTGATATGCAACAGAGAGACC 1220
Qy 380 -----LysTh 381
Db 1221 ATTGTCAAGCTTACGTCACGAATCCGTATTTCACGCGACGTGAATGATATGCTGCTC 1280
Qy 381 rLeuCybLysGlyAspArgGlnAlaProProGlyPro----- 393
Db 1281 ACTATCATCCGGGAGAAAGACGCGCCCTTAAATCCCTGCAACAGACAGACAGAGGTT 1340
Qy 394 -ProAlaArgPheProArg---ProLettPserAlaSerProProArgAlaProArgSe 412
Db 1341 CAATGCCCGTTGCTGAGACTCCCGACGTGCTCTCAGT---CGAACAGACAGCCGATC 1397
Qy 412 rSerThrProCybProGlyGlyAlaValAlaArgGluAspThrTyreProAlaGlyThrGlnG 432
Db 1398 CAAGATGCCCGCCGCTCC-GGAGGC---AGATACGAGGCTTACAGTCAAGCTATGA-GG 1452
Qy 432 yValPro-----SerProAlaLeuAlaGlnGlyGlyProGlnGlySerTr 447
Db 1453 TATTCACACACTTTTACGGAACCTGACCCCTGGCTGAAGTCGTCAGCGTGAGACATTTG 1512
Qy 447 pArgPheLeu-----GlnTPRAsnSerMetProArgLeuProThrAs 461
Db 1513 GCCAAAATATGTTTAAAGCTTACTGATTAAGAGGAACTGCTTCTGACTTCGACAGCC 1572
Qy 461 pLeuAspValGluGlyPro 467
Db 1573 TCATCAATGATGGGCT 1591

RESULT 13
US-09-270-767-9899/c
; Sequence 9899, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9899
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9899

Alignment Scores:
Pred. No.: 0.000168 Length: 806
Score: 194.50 Matches: 79
Percent Similarity: 41.1% Conservative: 44
Best Local Similarity: 26.4% Mismatches: 111
Query Match: 6.5% Indels: 67
DB: 3 Gaps: 9
US-10-071-838-2 (1-549) x US-09-270-767-9899 (1-806)
```

```
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 788 GCGTTCCAGAGACATCGCGTGAAGAAAGATCTGGCAGAACTGAGCCAGCTCGAGGCG--- 732
Qy 121 LysLeuLysAsnProGlyValArgTyrGlnIleMetLysGlnLysGlyLysArgSerArgL 140
Db 731 AGAATGGAATGAACGACAGATTCATTAATGACCAAGAAACCAAACTGTGAGACC 672
Qy 141 HisIleGlnAlaArgLysAspArgValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 671 GTTATCCAG-----CGGACATCATTCGCACTTTCGCCGCGCAAAATGCTTCAAA 621
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 620 GAATTTGGCGGTTCTGCGCCAAATGACATTTTAAAGTGTCCAAACGATGCGCTTCAT 561
Qy 181 AsnProGluValGlyTyrCybArgAspLysSerHisIleAlaAlaLeuPheLeuTyr 200
Db 560 GACAGCGAGGTTGATATGTCAGGGTCTAAGTTTCATAGCAGCTAGTCTGCTCTTCAT 501
Qy 201 LeuProGluGluAspAlaPheThrAlaLeuValGlnLeuLeuAlaserGluArgHisSer 220
Db 500 ATGCTGAGAGAGATGCGTTCTGTGTTCTGTAGCGCTTATGTAACGAC----- 453
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAsp----- 237
Db 452 -----TATGGCTCCGTATCTTCACAA 429
Qy 238 -----GlnGlnGluHisValAlaAlaThrSer 246
Db 428 GCCGATTCAGAGCTCTTACCTTCCTGCTCTACCAATGAGAGCAATGCAAGATCAG 369
Qy 247 GlnProLysThrMetCysLysHis-GlnAspLysLysAspLysCysGlyGlnCysSerPro 266
Db 368 CTGCCCACACTGACACAAACATTCACGCGCTGGCATGAGAACGACATGACGCTCC 309
Qy 266 uGlyCybLeuIleArgIleLeuIleAspGlyLysSerLeuGlyLeuThrLeuArgLeuTr 286
Db 308 CAATGG-TTCCTAACCTTATATACAGCGGAGATCCCATTTGTGTTGTTCCACGTCGT 250
Qy 286 pAspValTyrLeuValGlnGlyGlnAlaLeuMetProLettPheThrArgIleAlaPheL 306
Db 249 GGAATGTTTCTTACTGAGCGA-----CTACCTGTG----- 219
Qy 306 eValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTPRAlaArgPheCysAs 326
Db 218 -CTTTCAGGTGGCGGTTACTCTTATCAATCTGTGAATCC----- 177
Qy 326 nArgPheValAspThrTPRAlaArgAspGluAspThrValLeuLysHisLeuArgAlase 346
Db 176 -----GACTTCGCGCACTGATTTTGAAGGCACTTCAAGATATTTCCGGGTAA 127
Qy 346 rMetLysLysLeuThrArgLysGlnLysAspLeuProProAlaLysProGluGlnG 366
Db 126 GTTGCGAAAGAAAGTGGCGACGCTCC-----ACGCAAGG-ACGCAAGG 86
Qy 366 ySerSerAlaSer-----ArgProValProAlaSerArgGlyGly 379
Db 85 TGATGAAGCAAGCTGCGAAGCTAAGATAAAAACTGAAGCATGACGAGAGAGA 31

RESULT 14
US-08-363-300-1
; Sequence 1, Application US/08363300
; Patent No. 5700927
; GENERAL INFORMATION:
; APPLICANT: Zon, Leonard and Richardson, Paul
; TITLE OF INVENTION: Tbcd Gene and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
```

```

1 STATE: MA
2 COUNTRY: USA
3 ZIP: 02110-2804
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Patentin Releasee #1.0, Version #1.30B
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/08/363,300
12 FILING DATE:
13 CLASSIFICATION: 514
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Clark, Paul T.
16 REGISTRATION NUMBER: 30,162
17 REFERENCE/DOCKET NUMBER: 04590/009001
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: (617) 542-5070
20 TELEFAX: (617) 542-8906
21 INFORMATION FOR SEQ ID NO: 1:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 4039 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: double
26 TOPOLOGY: linear
27 MOLECULE TYPE: DNA (genomic)
28 FEATURE:
29 NAME/KEY: CDS
30 LOCATION: 403..3829
31
32 US-08-363-300-1

```

```

Db      3698 -----ATGTTGAGAGAGCTGCGAGGCAAAAGCCGCCAGCACT 3739
QY      512 AapSerAapGlnGlyThrProPheArgAlaArgAapGlnGlnPro 526
Db      3740 CCAGAGCCAGACTGACCCAGCTGGAGCCAGCCGATTGACCG 3784

RESULT 15
US-09-762-311-3
/ Sequence 3, Application US/09762311
/ Patent No. 6825004
/ GENERAL INFORMATION:
/ APPLICANT: BLUMENFELD, Maria
/ APPLICANT: BOUGUELERET, Lydie
/ APPLICANT: CHUMAKOV, Ilya
/ TITLE OF INVENTION: Nucleic Acids Encoding Human TBC-1 Protein And Polymorphic Marker
/ FILE REFERENCE: 46.US2.PCT
/ CURRENT APPLICATION NUMBER: US/09/762,311
/ PRIOR FILING DATE: 2001-06-21
/ PRIOR APPLICATION NUMBER: US 60/095,653
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: Patent.pm
/ SEQ ID NO 3
/ LENGTH: 3983
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 171..3725
/ NAME/KEY: polyA signal
/ LOCATION: 3942..3947
/ OTHER INFORMATION: AATAAA
/ NAME/KEY: misc_feature
/ LOCATION: 36
/ OTHER INFORMATION: n=a, g, c or t
US-09-762-311-3

Alignment Scores:
Pred. No.: 0.0258 Length: 3983
Score: 175.50 Matches: 59
Percent Similarity: 42.2% Conservative: 41
Best Local Similarity: 24.9% Mismatches: 114
Query Match: 5.9% Indels: 23
DB: 3 Gaps: 4

US-10-071-838-2 (1-549) x US-09-762-311-3 (1-3983)
QY      85 TTPGLuLyEYrLySserSer-----ArgLyELeu 94
Db      2490 TGGGAAAAGATGCTTACCACTCCAGAAAGATCAAAAATTAAATTGACATGAAATAATG 2549
QY      95 ILeaAparGAlaTyLySgLyMeLProMeLAsnIleArgLyProMeLTPSerValLeu 114
Db      2550 CACTCGGCTGTGGGCAAGGTGTGCCACGTATCACCGAGGTGAATCTGAAATTCTA 2609
QY      115 LeuAenIleGluGluMetLyELeuLyAsnProGlyArgTyGln-----Ile 130
Db      2610 GCTGAGCAATTCACCTTAAACACAGTTTCCAGCAACAGACAGCCAAAGATGTGCCA 2669
QY      131 MetLySgLySgLyLySAsrSerSerGluHisIleGlnArgIleAparGAspValSer 150
Db      2670 TACAAAGAACTCTTAAAGCAGCTGACTTCCAGACAGATCGATTCTTATGACCTTGGG 2729
QY      151 GlyThrLeuArgLyHisIlePhePheArgAparGlyTyThrLySgIleArgGluLeu 170
Db      2730 CGAACCTTTCCTACACCACTACTCTCTGCCAGCTGGAGACAGACATATGCTT 2789
QY      171 LeuHisIleLeuLeuAlaTyGluGluTyArgAsnProGluValGlyTyCyAparGAspLeu 190
Db      2790 TACAAACATTTTGAAGGCTACTCACTTCTAGACCAAGAGAGTGGATATTGCAAGTCTC 2849
```

```

QY      191 SerHisIleAlaIleuPheLeuLeuTyLyLeuProGluGluAspAlaPheTrpAlaLeu 210
Db      2850 AGCTTTGTAGAGGCAATTTTGCTTCTTCAATGAGTGAAGAGGCGTTTAAATGCTC 2909
QY      211 ValGluLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAenGlyGly 230
Db      2910 AAGTTCTGATGTTTGAC-----ATGGGGCTGGGAAACAGTATCGGCCAGCATGATT 2963
QY      231 ThrValGlnGlyLeuGlnAspGlnGlnHisValIleAlaThrSerGlnProLySthr 250
Db      2964 ATTTACAGATCCAGATGTACAGCTCTGAGGTTCTTCATGATTACACAGAGACCTC 3023
QY      251 MetGlyHisGlnAspLySAspLeuCySgLyGlnCySerProLeuGlyCySLeuIle 270
Db      3024 TACAATCACCTGAGAGAGACAGATCGGCCCCAGCCTGACGCTGCCCTGTTCTTC 3083
QY      271 ArgIleLeuIleAparGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAparValTyLeu 290
Db      3084 ACCATGTTTGCTTCACAGTTCCCGCTGGGATTCGTAGCCAGAGTCTTGATATGATTTT 3143
QY      291 ValGluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLySVal 307
Db      3144 CTTCAGGA-----ACAGAGTCAATTTTAAATG 3173
```

Search completed: April 5, 2006, 14:54:45
Job time : 305 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:42:42 ; Search time 6211 Seconds

(without alignments)
4135.583 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWAQEREDII.....TSGPCGCLHLESQPPPGF 549

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=x1h
-Q=/abs/ABSSWB.epool/US10071838/runac_05042006_141915_5173/app_query.faeta_1
-DB=EST -QFMT=faeta -SUFFIX=xtc -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -ICPU=3
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pro -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808
-USER=US10071838 -CGEN_1_1_5315 -runac_05042006_141915_5173 -NCPU=6 -ICPU=3
-NO MMWP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hic: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_ges1: *
10: gb_ges2: *
11: gb_ges3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match	Length	DB ID	Description
1	2963	99.3	2072	4 HSM801828	AL136860 Homo sapi
2	2820	94.5	2497	4 BC047739	BC047739 Homo sapi
3	1513	50.7	915	5 BQ923700	BQ923700 AGENCOURT
4	1455	48.8	954	5 BQ526609	BQ526609 AGENCOURT
5	1440	48.3	909	3 BI753688	BI753688 603023590
6	1397.5	46.8	868	6 CD653466	CD653466 AGENCOURT
7	1342	45.0	881	5 BU542901	BU542901 AGENCOURT

8	1338.5	44.9	1013	3 BM560320	BM560320 AGENCOURT
9	1335	44.8	919	3 B1522600	B1522600 603175612
10	1324	44.4	930	5 BQ920917	BQ920917 AGENCOURT
11	1323	44.4	1141	3 BM553146	BM553146 AGENCOURT
12	1302.5	43.7	1079	3 BM563571	BM563571 AGENCOURT
13	1272.5	42.7	831	3 B1828272	B1828272 603078163
14	1243	41.7	709	2 B1088323	B1088323 602852793
15	1240.5	41.6	946	5 BQ891586	BQ891586 AGENCOURT
16	1211	40.6	3109	4 BC037244	BC037244 Homo sapi
17	1181.5	39.6	839	6 CA487310	CA487310 AGENCOURT
18	1181	39.6	864	3 BM914524	BM914524 AGENCOURT
19	1171.5	39.3	840	3 B1829764	B1829764 603079835
20	1168	39.2	667	3 B1827136	B1827136 603077337
21	1162.5	39.0	882	3 B1520522	B1520522 603071531
22	1162	39.0	634	7 CN262008	CN262008 170005377
23	1138	38.1	839	3 B1601698	B1601698 603249940
24	1118	37.5	829	3 B1597531	B1597531 603247059
25	1117	37.4	618	3 BM843014	BM843014 K-EST0120
26	1117	37.4	630	3 BM842553	BM842553 K-EST0119
27	1117	37.4	727	1 AL042933	AL042933 DKFZp334K
28	1116.5	37.4	848	5 BQ942166	BQ942166 AGENCOURT
29	1113	37.3	628	3 BM762465	BM762465 K-EST0043
30	1112	37.3	742	3 B1832051	B1832051 603075809
31	1111	37.2	986	5 BX346181	BX346181 BX346181
32	1097	36.8	628	8 DN999748	DN999748 TC104423
33	1067	35.8	635	1 AV650439	AV650439 AV650439
34	1065	35.7	633	8 DR006725	DR006725 TC105000
35	1056	35.4	802	7 CR766757	CR766757 DKFZp468D
36	1049	35.2	650	3 B1830812	B1830812 603080926
37	1026.5	34.4	979	3 B1830468	B1830468 603073339
38	994	33.3	573	3 BM842932	BM842932 K-EST0120
39	950	31.8	539	5 BX281608	BX281608 BX281608
40	932	31.2	545	7 CN262009	CN262009 170006001
41	931.5	31.2	1045	3 BM563007	BM563007 AGENCOURT
42	925	31.0	944	5 BQ709135	BQ709135 AGENCOURT
43	904	30.3	498	5 BX279913	BX279913 BX279913
44	891	29.6	671	3 B1828366	B1828366 603079333
45	887.5	29.8	875	9 A0788033	A0788033 HS_3100_A

ALIGNMENTS

RESULT 1
HSM801828
LOCUS HSM801828 2072 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA, CDNA DKFZp434P2235 (from clone DKFZp434P2235).
ACCESSION AL136860
VERSION AL136860.1 GI:12053224
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 2072)
Fousterka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstadtstr Landstr.1, D-85764
Neuerberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center).
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp434P2235) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/ci.cgi?CID=DKFZp434P2235
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/
Location/Qualifiers

REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
COMMENT
FEATURES

```
source
1. .2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434P2235"
/db_xref="taxon:9606"
/clone="DKFZp434P2235"
/clone_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
/notice="TBC1 domain family, member 3"
1. .2072
/gene="DKFZp434P2235"
99. .1748
/gene="DKFZp434P2235"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB6794.1"
/db_xref="GI:12053225"
/db_xref="UniProt/Swiss-Prot:Q81ZP1"
/translation="MDVYEVASGMAQGERDITMKYEGHRAGLPEDEKPKPFPSYNN
NVDHGLVHETELPELTAREAKQIRREISRSKWDMDGDEKTSKRLIDRAYKGM
PMNIRGPMMSVLNTEEMKLNPKRYQIMKSGKSSBHIQIRIDVSGTLRKHIFER
DRYGTQREBLHILAYEYNEVGYCDLSHIALFLYLPEDAFWALQULASER
HSLQGFHSPNGSTVOGLDOOEHVAVATOPKTMGHODKDLGGCCSLRLILIDG
ISLGLTLRIMPVLYEGSOALMPITRIAPKQOKRITKTSRGPMARCNFVDTPAR
DEDTVLKLRASMKLTKRKGDLPAPAPBQSSASRPVPSRGKTKCKGDRAPPC
PPAPPRPISASPPRPRSSPCPGAVRBDTPVGTQVSPALAGGQDSMRPL
QMSMRPLRDLDEGPMFRHYDFRQSCWVRASIQEDQLAPCQWQAEHRAERSAFA
PSTDSDQCTPPRARDQCAPTSQCLGLHLSQPPGF"
```

Alignment Scores:

Pred. No.:	1.28e-194	Length:	2072
Score:	2963.00	Matches:	545
Percent Similarity:	99.64	Conservative:	2
Best Local Similarity:	99.34	Mismatches:	2
Query Match:	99.34	Indels:	0
DB:	4	Gaps:	0

US-10-071-838-2 (1-549) x HSN801828 (1-2072)

```
QY 1 MetAEPValValIGluValAlaGlySerTPRPAIAGInGluArgGluAspIleIleMet 20
DB 99 ATGAGCGTGTAGAGAGTCGCGGCGAGTGTGTGGGCAAGAGCGAGAGCATCTTATG 158
QY 21 LysTYrGluYsGlyYHleArgAlaGlyLeuProGluAspLysGlyProLysPhePheArg 40
DB 159 AAATACGAAAAGGAGCACCGAGCTGGCTGCGCAGAGGACCAAGGGGCTTAAGCCTTTTGA 218
QY 41 SerTYrAenAenAenValAenHleValIleValHleGluThGluLeuProPheLeu 60
DB 219 AGCTACCAACCAACAGCTGCATCTTGGGATTTGACATGAGACCGAGCTGCTCTCTG 278
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAen 80
DB 279 ACTGCGCGGAGCGGACCAAGCAATTTCGCGCGGAGATCAGCCGAAAGCAAGGAGTGAAT 338
QY 81 MetLeuGlyAspTrpGluYsTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100
DB 339 ATGCTGGGAGACTGGGAGAAATATACAAAGCAGACAAAGCTCATGATCAGACGTACAG 398
QY 101 GlyMetProMetAenIleArgGlyProMetTrpSerValLeuLeuAenIleGluGluMet 120
DB 399 GGAATGCCCATGAAATCATCCGCGGCGCGAGTGTGTCATCTCTCAACCTAGGAAATG 458
QY 121 LysLeuLysAenProGlyYArgTYrGlnIleMetLysGluYsGlyLysArgSerGlu 140
DB 459 AAGTTGAANAACCCCGAAGATACCAAGATCATGAAGAGAGAGGCAAGAAAGCATCTGAG 518
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 519 CACATCCAGCGCATGACCGGAGCGTAAGCGGAGCATTTAAGAAACATATATCTTCAAG 578
```

```
QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
DB 579 GATGATATCCGAACCAAGACAGCGGGAACTACTCCACATCTCTCGCATATGAGAGATAC 638
QY 181 AsnProGluValGlyTYrCyArgAspLeuSerHisIleAlaAlaLeuPheLeuTYr 200
DB 639 AACCAGAGGAGGCTACTGACAGGACCTGAGCCACATCCCGCTTGTCTCTCTAT 698
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 699 CTTCCTAGAGAGATGATCTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGAGACTCC 758
QY 221 LeuGlnGlyPheHisSerProAenGlyYThrValGlnIleLeuGlnAenArgGlnGln 240
DB 759 CTGACGAGATTTTCACACCCCAATAGCGGAGACCTCCACAGGGGCTCCAAAGCAACAGAG 818
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHleGlnAspLysValAspLeuCy 260
DB 819 CATGTGTAGCGACGTCAACAACCAAGACCATGGGCGATCAGGACAAAGAAATCTATGT 878
QY 261 GlyGlnCySerProLeuGlyCyLeuIleArgIleLeuIleLeuIleLeuIleSerLeuGly 280
DB 879 GGGCAGGTTCCTCCGTTAGGCTGCTCATCCGATATGTATGACGGGATCTCTCGGG 938
QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGluIleAlaLeuMetProIle 300
DB 939 CTCACCTCGCGCTGTGGAGCTGTATCTGTGTAGAACGCAACAGGCGTGTATCCGATA 998
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCyArgIle 320
DB 999 ACAAGATTCGCTTTAAGTTACAGCAGAGCGCTCCAGAAAGCTCCAGGTGTGGCCCG 1058
QY 321 TrpAlaArgPheCyAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1059 TGGCAGAGTGTTCGAAACCGGTTGTTGATACCTGGGCGCAGGAGTGAAGACATGTCTC 1118
QY 341 LysHisLeuArgAlaSerMetLysLeuLeuThrArgLysGlnIleAspLeuProPro 360
DB 1119 AAGCATCTTAGGCGCTCTTAGAAGAACTAAGAAAGAAAGAGGAGCATGCCACCCCA 1178
QY 361 AlaLysProGluGlnIleYSerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1179 GCCAAACCCGAGCAAGGAGTGTGCGGCATCCAGGCGCTGTGCGGCTTACGTTGCGGAG 1238
QY 381 ThrLeuCyLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg 400
DB 1239 ACCCTTCGCAAGGGGAGCAGGACAGCGCCCTTCAGAGCCACCAAGCCGAGTCCCGGACC 1298
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerThrProCyProGlyGlyAla 420
DB 1299 ATTTGGTCAGCTTCCCGCCAGCGGACCTGCTCTTCCACACCTGTCTCGTGGGGCT 1358
QY 421 ValArgLysPheThrTYrProValGlyYThrGlnIleYValProSerProAlaLeuAlaGln 440
DB 1359 GTCCGGGAAAGCACCTACCTGCGGACATCAGAGGTGTGCCAGCCCGGCTGTGCTCAG 1418
QY 441 GlyGlyProGlnIleYSerTrpArgPheLeuGlnIleTrpAsnSerMetProArgLeuProThr 460
DB 1419 GGAGAGCTTCAGGGGTCTCTGAGATTCCTGAGTGAATCTCAGAGTGAATCCAGTCCCGCTCCAAAG 1478
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTYrAspPheArgGlnSerCySerTrpVal 480
DB 1479 GACCTGAGCGTAGAGGCGCTTGTGTTCGCAATTAATTTCAAGACAGAGTGTCTGGGTC 1538
QY 481 ArgAlaIleSerGlnIleAspGlnLeuAlaProCyTrpGlnAlaGluHisProAlaGlu 500
DB 1539 CGTGCATATCCAGAGAGACCACTGAGCCCTGTGCGAGGCTGAAACCTGTGCGGAG 1598
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrArgSerArgGlnIleYThrProPheArg 520
DB 1599 CGGGTGAATGTGGCTTTCGCTGCAACGACGACATGATTCGACGAGGAGCACCCTTCAAG 1658
```


QY 521 A1a9a9p9l9u9i9n9p9r9o9y9a9l9a9p9r9o9t9h9s9e9r9g9l9y9p9r9o9c9y9s9l9e9u9h9i9l9e9u9 540
 Db 1659 GCTAGGAGCAACAGCAGTGTGCTCCCACTCCAGGGCTTCCCTCTGCGGCTTCACTTG 1718

QY 541 GluSerSerGlnPheProGlyPhe 549
 Db 1719 GAAAGTCTCAGTTCCTCCAGGCTTC 1745

RESULT 2
 BC047739 2497 bp mRNA linear HTC 20-MAR-2003
 LOCUS Homo sapiens, similar to Rab GTPase-activating protein PRC17, clone
 DEFINITION IMAGE:5743752, mRNA.
 BC047739
 ACCESSION BC047739.1 GI:29126829
 VERSION
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2497)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAR-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 CONTACT: MGC help desk
 Email: g9abs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www.hhg.scf.stanford.edu
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

REMARK
 COMMENT
 Clon distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
 Series: IRAC Plate: 98 Row: d Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 14149984
 This clone has the following problem: retained intron.
 FEATURES
 source
 1..2497
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5743752"
 /cissue_type="Brain, adult medulla"
 /clone_lib="NIH_MGC_119"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORTS"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,32e-184 Length: 2497
 Score: 2820.00 Matches: 542
 Percent Similarity: 79.0% Conservative: 2
 Best Local Similarity: 78.7% Mismatches: 5
 Query Match: 94.5% Indels: 140
 DB: 4 Gaps: 2

US-10-071-838-2 (1-549) x BC047739 (1-2497)

QY 1 MetA9pValI9u9Val9a9l9a9g9l9s9e9r9t9r9p9a9l9g9l9u9a9r9g9l9u9a9p9l9e9l9e9m9e9t9 20
 Db 89 ATGACGTGTAGTAGGTGCGGGCAGTGTGGTGCGCAAGAGCGAGGAGCATCATTAAG 148

QY 21 lyeTgTglu9y9e9g9l9h9i9a9r9a9l9a9g9l9e9u9p9r9o9l9u9a9p9r9o9p9h9e9a9r9g9 40
 Db 149 AAATACAAAGAGGACACCAAGCTGGGCTCCAGAGGACAAAGGGCTTAAGCTTTTCGA 208

QY 41 SerTyra9a9a9a9a9Val9a9Phe9h9i9le9u9l9e9a9l9 52
 Db 209 AGCTACAAACAAACAGTCGATCATTTGGGGATTGTACAGTCGGGTCAACCTCTTTCAG 268

QY 52 52
 Db 269 GACAGAACTCTCTCTGACTCCCTGACGTCCAGCCGAGGTGTAGGCCAGAGTGC 328

QY 52 52
 Db 329 TGGGGCCCATCTAGGAGCCGGTGGGAATGAGACTGGGTCAGGCCCTGGGGCC 388

QY 52 52
 Db 389 TCAGCAGTCTGTGCGCAAGTGAACAAAGAGAGCGGGCAGCCTGAGGGTCTGGCCCT 448

QY 53 53
 Db 449 GTCTACTTGGAGACAACCCCGTGAACGGAGCTGCTCTGACTGCGGGAGCGGAA 508

QY 66 66
 Db 509 GCAAAATTCGGCGGAGATGACCGCAAGACCAAGTGGGTGATATGCTGGAGACTGGGA 568

QY 86 86
 Db 569 GAATACAAAGAGAGCAAGAAAGCTCATAGTCAAGCTTCAAGGAAATGCCATGAACAT 628

QY 106 106
 Db 629 CCGGGCCCGATGTGTGCTGCTCTGCAACATTGAGAAATGAAGTTGAAAAACCCCG 688

QY 126 126
 Db 689 AAGATACCAAGATCATAGAGAGAAAGGCAAGAGTATGACACATCCAGGGCATCGA 748

QY 146 146
 Db 749 CCGGAGCAGTAAGGGGACATTAAGAGACATATATCTTCAAGGATCGATCGGAACCA 808

QY 166 166
 Db 809 GCAGCGGAACTACTCCATCTCTGCGCATATGAGAGTACAAACCCGAGGTGGCTA 868

QY 186 186
 Db 869 CTGACAGGACCTGAGCACATCGCGCTTGTCTCTTATCTTCTGAGAGGATGC 928

QY 206 206
 Db 929 ATTCTGGGACCTGTGCAAGCTGCTGCGCATGAGAGCACTCCCTCAGGAGATTTACAG 988

QY 226 226
 Db 989 CCCAAATGCGGGAGCGCTCAGGGGCTCCAAAGCAACAGAGCATGTGTACCAAGTC 1048

QY 246 246
 Db 1049 ACAACCCAAAGACATGCGGACATCGAGCAAGAAAGTCTATGTGGCAGGTGCCCGTT 1108

QY 266 266
 Db 1109 AGGCTGCTCATTCGGATATTAATGACCGGATCTCTCGGGCTACCCCTGGCCTGTG 1168

QY 286 286
 Db 1169 GACGCTGTATGTGTAGAGGCAAGCGCGCTGATCGGATACAAAGATGCGCTTTAA 1228

QY 306 306
 Db 306 sValGlnGln----- 309

```

Db      1229 GGTTCACAGACAGTAAGTCTACGTGTGCCACAGGGGCTCTGGGAGACCTCTGGGGTCAAGACC 1288
          |||||
Qy      309 -----
Db      1289 CCGACTGGCCCGAGGGGAGCTTCTCTACACTGTCTCATGTATCTCTGTCTGGCCAGAGA 1348
Qy      309 -----
Db      1349 GGGAGGCTTGGCCAGGTGGGCTGGGAGAGACCTGTGACACCGAGCCATCCCCACATG 1408
Qy      310 -----
Db      1409 ACCCAGATGAAAGTCGAGAGTGTGAGCCCTTCACGAAACGCTCCAGAGTGTGGCCGTGGG 1468
          :|:|:|
Qy      322 laarpheCyasnaarphveValaaphrThrTpaLaarGAspGluAaphThrValleuLysH 342
          CACGTTTTCACACCGGTTCCGTGTATCTGTGGCCAGGAGATGAGACACTGTGTCAAGC 1469
          |||||
Qy      342 isleuAargAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProProAla 362
          |||||
Db      1529 ATCTTAGGGCTCTATGAAAGAACTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1588
          |||||
Qy      362 YSPROGLINGlySerSerAlaSerArgProValProAlaSerArgLysGlyLysThrL 382
          |||||
Db      1589 AACCAGAGCAAGGCTCTCGGCATCCAGGCTGTGCGGCTTCACGTGGGGAGAGACC 1648
          |||||
Qy      382 euCyalyeGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgProLett 402
          |||||
Db      1649 TCTCAAGAGGGAGACAGGAGGCGCTCCAGGCGCCACAGCCGCGTTCGCGGGCCCATTT 1708
          |||||
Qy      402 rPserAlaSerProProArgAlaProArgSerSerThrProCyAspProGlyGlyAlaVala 422
          |||||
Db      1709 GGTACGCTTCCTGTCACAGGAGACTCTGTTTCCACACCTGTCTGTGTGGAGCTGTCTCC 1768
          |||||
Qy      422 rGgluAaphThrTyxProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGly 442
          |||||
Db      1769 GGGAGAGACACTTACCTGTGGGCACTCAGGGGTGTCCAGCCGCGCTGTGTCAGGGAG 1828
          |||||
Qy      442 lYProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProTrpArgL 462
          |||||
Db      1829 GACCTCAGGGTTCCTGGAGATTCCTGCAGTGGACCTCATGCCCCGCTCCCAACGAGACC 1888
          |||||
Qy      462 euAAspValGlnGlyProTrpPheArgHsTyxAspPheArgGlnSerCysTrpValArgA 482
          |||||
Db      1889 TGGAGTGAAGAGGCGCTTGTTCCGTCATATGATTTCAAGACAGAGCTGTGGCCCTGTG 1948
          |||||
Qy      482 laileSerGlnGluAspGlnLeuAlaProCySTrPGLnaGlnHsIspProAlaGluArgV 502
          |||||
Db      1949 CCATTTACCAAGAGAGACCAAGCTGGCCCTGCTGGCAAGCTGAACACCCTGGGAGCGGG 2008
          |||||
Qy      502 aIArgSerAlaPheAlaAlaProSerThrAspSerArgGlnGlyThrProPheArgAla 522
          |||||
Db      2009 TGAGATCGGCTTCCCTGCACCCAGCACTGATTCGAGCCAGGCAAGCCCTTCACAGCTA 2068
          |||||
Qy      522 rGaAspGlnGlnProCyAspAlaProTrpSerGlyProCyLeuCyAspGlyLeuHsIleuGln 542
          |||||
Db      2069 GGGAGGAACAGCAGTGTCTCCACCTCAGGGGCTTGCTCTGGGCTTCACCTTGAGAA 2128
          |||||
Qy      542 eSerGlnPheProProGlyPhe 549
          |||||
Db      2129 GTTCTCAGTTCCCTCAGGCTTC 2151
          |||||

```

```

RESULT 3
BQ923700      915 bp      mRNA      linear      EST 20-AUG-2002
LOCUS        BQ923700
DEFINITION   AGENCOURT 8798484 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6425622
ACCESSION    BQ923700
VERSION      BQ923700.1 GI:22338731
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 915)
NIH-MGC <http://mgs.nci.nih.gov/>,
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
Plate: LUC2607 row: 1 column: 07
High quality sequence stop: 639.
Location/Qualifiers

FEATURES

source

```

1..915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6425622"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 101"
/notes="Organ: lung; Vector: pCMV7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

```

ORIGIN

Alignment Scores:

```

Pred. No.:      1,13e-94      Length:      915
Score:          1513.00      Matches:      295
Percent Similarity: 94.5%      Conservative: 2
Best Local Similarity: 93.9%      Mismatches: 6
Query Match:    50.7%      Indels:      12
DB:              5      Gaps:      1

```

US-10-071-838-2 (1-549) x BQ923700 (1-915)

```

Qy      137 ArgSerSerGlnHsIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHs 156
          :|:|:|
Db      1 AAGGCATCTGAGCAATCCAGCATCGACCGGAGCGTAAAGCGGAGCACTTAAGAGAGCAT 60
          |||||
Qy      157 lIlePhePheArgAspArgTyxGlyThrTyxGlnArgGlnArgGlnLeuLeuHsIleLeuLeuAla 176
          |||||
Db      61 ATATTCTTCAAGAGAGTCGATACAGAACCAAGAGAGGAGAACTTCTCCATCTCTCTGGCA 120
          |||||
Qy      177 TyxGlnGlnTyxAspProGluValGlyTyxGlyAspLeuSerHsIleAlaLeu 196
          |||||
Db      121 TATGAGAGTATTAACCCGAGAGTGGGCTATCGAGGAGCTGAGCAATGCGGCTTGG 180
          |||||
Qy      197 PheLeuLeuTyxLeuProGlnGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSer 216
          |||||
Db      181 TTCTCTCTATCTTCTCTGAGAGAGATGCAATTCGGGCACTGTGCACACTGTGCGCAGT 240
          |||||
Qy      217 GluArgHsIleSerLeuGlnGlyPheHsIleSerProAsnGlyGlyThrValGlnGlyLeuGln 236
          |||||
Db      241 GAGAGGCACTCCCTGCAGAGGATTTCAAGCCCAATGCGGAGCCGCTCAGAGGCTCCAA 300
          |||||
Qy      237 AspGlnGlnGlnHsValaValaAlaThrSerGlnProLysTrpMetGlyHsIleGlnAspLys 256
          |||||
Db      301 GACCAACAGAGACATGTGTAGCCGTCACACCAACCAACATGAGGAGATCAGAGCAAG 360
          |||||
Qy      257 LysAspLeuCyAspGlyGlnCySerProLeuGlyCyLeuLleArgIleLeuLleAspGly 276
          |||||

```

Db 361 AAGATCATATGTCGACATGTTCCCGTAGAGCTGCTCATCCGATATTGATGACGGG 420
Qy 277 ILeSerLeuGlyLeuThrLeuArgLeuTrpAspValIyrlleuValGluGluGlnAla 296
Db 421 ATCTCTCGGGCTCACCTGCGCTGTGGACCTGTATCTGTAGAGGGAACAGCGG 480
Qy 297 LeuMetProIleThrArgIleAlaPheIleValGlnGlnIlySerArgLeuThrIleThrSer 316
Db 481 TTGATGCCATTAACAAGAAATCGCTTTAAGTTACAGCAAGAGCCCTCAGAAAGATGCC 540
Qy 317 ArgCysGlyProTrpAlaArgPheCysAsnArgPheValIlePheTrpAlaArgAspGlu 336
Db 541 AGGTGTGGCCGTGGGACGTTTTCACACCGGTTCGTATACCTGGGCGAGGATGAG 600
Qy 337 AspThrValLeuIlySerHisIleuArgAlaSerMetIlySerLeuThrArgIlyGlnIlyAsp 356
Db 601 GACACGTGTCTCAAGACATCTTANGGCTCTATGAAAGAACTAACAAAGAGCAGGAGAC 660
Qy 357 LeuProProProAlaIlyProGluGlnIlySerSerAlaSerArgProValProAlaSer 376
Db 661 CTCACACCCCAAGCCAAACCCGAGCAAGGCTGTGCGCATCCAGGCTGTGCGGCTTCA 720
Qy 377 ArgGlyIlyIyThrLeuCysIyGlyAspArgGlnAlaProProGlyProProAlaArg 396
Db 721 CGTGGCGGAAAGACCTCTGCAAGGGGGAAGGAGGCCCTCCAGC-CCACCAAGCCCGG 779
Qy 397 PheProArgProIleTrpSerAlaSerProProArgAlaProArgSerSerThrProCys 416
Db 780 TTCCTGGGGGCCATTGGTACAGCTTCCCGCACAGGACCTTCATCTCCACACCTCTGT 839
Qy 417 -ProGlyIlyAlaValAlaArg-GlyuSerThrIy- ProValGlyThrGlnIlyValProSer 435
Db 840 TCTGTGGGGGCTGTCCGGGGAAGACACCTACCTCTGGGGGCACTCAAGGTGTGCCAN 899
Qy 436 ProAlaLeuAlaGlnIlyGlyProGlnIlySerTrp 447
Db 900 CCC-----GGGCCCTCGG 911

RESULT 4
BUS26609 954 bp mRNA linear EST 13-SEP-2002
LOCUS
DEFINITION
AGENCOURT 10181753 NIH MGC 101 Homo sapiens cDNA clone
IMAGE:5636307 5', mRNA sequence.
BUS26609
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 954)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csabds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2697 row: h column: 03
High quality sequence stop: 627.
Location/Qualifiers
1..954
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5636307"
/issue_type="epidermoid carcinoma, cell line"

Alignment Scores:
Pred. No.: 1.24e-90 Length: 954
Score: 1455.00 Matches: 283
Percent Similarity: 94.4% Conservative: 2
Best Local Similarity: 93.7% Mismatches: 11
Query Match: 48.8% Indels: 6
DB: 5 Gaps: 2

US-10-071-838-2 (1-549) x BUS26609 (1-954)

Qy 173 ILeLeuLeuAlaTrpGluGluIyTrpAspProGluValGlyIyCysArgAspLeuSerHis 192
Db 1 ATCTCTCGGGCAATGAGAGTATTAACCCGAGGTGGCTTACGACGACCTGAGCCAC 60
Qy 193 ILeAlaAlaLeuPheLeuLeuIyrlleuProGluGlnAspAlaPheTrpAlaValGln 212
Db 61 ATGCGCGCTTGTCT 120
Qy 213 LeuLeuAlaSerGluArgHisSerLeuGlnIlyPheHisSerProAsnGlyIyThrVal 232
Db 121 CTCTGTGGCAAGTAAAGGCACTCTGCAAGGAAATTCACAGCCAAATGCGGAGCCGTC 180
Qy 233 GlnGlyLeuGlnAspGlnGlnIlyHisValAlaIleThrSerGlnProIlySerMetGly 252
Db 181 CAGGGGCTCCAAACCAACCAAGAGCATGTGTGACCACTGACCAACCAAGACCATGGG 240
Qy 253 HisGlnAspIlyIyAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIle 272
Db 241 CATCAGGACAAAGAAAGTATGTGTGGGAGTGTTCCTCCGTTAGGCTCATCCGGATA 300
Qy 273 LeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValIyrlleuValGlu 292
Db 301 TTGATTGACGGAGATCT 360
Qy 293 GlyIlyGlnAlaLeuMetProIleThrArgIleAlaPheIleValGlnGlnIlySerArgLeu 312
Db 361 GCGGAACAGGCTGTGATGCGATTAACAAGATGCTTTAAGTTACAGCAAGAGCCCTC 420
Qy 313 ThrIyThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValIlePheTrp 332
Db 421 ACAGAAACGTCACAGGTGTGGCCGTGGGACGTTTTCACACCGGTTGATGATACCTGG 480
Qy 333 AlaArgAspGlnAspThrValLeuIlySerHisIleuArgAlaSerMetIlySerLeuThrArg 352
Db 481 GCCAGGATGAGACCTGTGCTCAAGACATCTTAAGGCTCTATGAAAGAACTAACAAAGA 540
Qy 353 IyGlnIlyAspLeuProProProAlaIlyProGluGlnIlySerSerAlaSerArgPro 372
Db 541 AAGCAGGGGAGCTGCAACCCCAAGCAACCCGAGCAAGGTCGTGCGATCCAGGCTT 600
Qy 373 ValProAlaSerArgIlyGlyIyThrLeuCysIyGlyAspArgGlnAlaProProGly 392
Db 601 GTGCCGCTTCAAGTGGCGGAAAGACCTCTGCAAGGGGGAAGAGGAGCCCTCCAGGC 660
Qy 393 ProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaProArgSer 412
Db 661 CACACAGCCCGGGTCCCGGGCCCATTTGGTACGTTCCCGCAGGGGACCTCGTTCT 720
Qy 413 SerThrProCysProGlyIlyAlaValAlaArgGluAspThrIy- ProValGlyThrGlnIly 432
Db 721 TTCACACCTGTCTGTGTGGGCTGTCCGGGAAACCACTTACCTGTGTGGCACTGAGGT 780

QY 433 ValProSerProAla-LeuAlaGlnGly-ArgGlnGlySerThrPheLeu--GI 451
DB 781 GTGCCACCGCCGCTGCTCAAGAGAACTCAGGGATCTCGGAG 840
QY 451 ntrpAsnSerMet---ProArgLeuProThrAspLeuAspVal---GlnGlyProTrp 468
DB 841 GGGAAACTCCATGGCCCCCGCTCCCAACGAGCTGGAGCTTAAAGGGGCCCTGG 898
RESULT 5
B1753688
LOCUS
DEFINITION 603023590P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193959 5',
mRNA sequence.
ACCESSION B1753688
VERSION B1753688.1 GI:15745266
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11485 row: d column: 24
High quality sequence stop: 884.
Location/Qualifiers
1..909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5193959"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1.27e-89 Length: 909
Score: 1440.00 Matches: 288
Percent Similarity: 94.8% Conservative: 3
Best Local Similarity: 93.8% Mismatches: 9
Query Match: 48.3% Indels: 7
DB: 3 Gaps: 0
US-10-071-838-2 (1-549) x B1753688 (1-909)
QY 56 GlnuPProPheLeuThrAlaArgGlnAlaLysGlnIleArgArgGlnIleSerArgLys 75
DB 3 GAGCTGCTCTCTGACTGCGCGGAGGCGAATTCGCGGAGATCAGCCGAGCA 62
QY 76 -SerLysTrpValAspMetLeuGlyAspTrpGlnLysTrpLysSerSerArgLysLeuI 95
DB 63 GAGCAAGTGGCTGGATATGCTCGGAGACTCGGGAATTCAAAGCAGCAGAAAGCTCAT 122

QY 95 eAspArgAlaTyTrpGlyMetProMetAsnIleArgGlyProMetTrpSerValLeu 115
DB 123 AGATCGAGCGGTACAGGGGAATGCCCATGAACTATCGGGGCCGATGTGTCAGTCTCT 182
QY 115 uAsnIleGlnGlnMetLysLeuLysAsnProGlyAspTyTrpGlnIleMetLysGlyVal 135
DB 183 GAACACTGAGGAATGAAATTGAAAAAACCCTGGAGATACAGATCTCATGAAGGAAAGG 242
QY 135 YLysArgSerSerGlnIleIleGlnArgIleAspArgAspValSerGlyThrLeuArgLys 155
DB 243 CAAGAGGTCACTGAGCACTACGCCATTCAGCCGAGAGCTTAAAGCGGACATTAAAGAA 302
QY 155 sHisIlePhePheArgAspArgTyTrpGlyTrpLysGlnArgGlnLeuLeuHisIleLeu 175
DB 303 GCATATATCTCTTACGGGATCGATACGGAACCAAGCAGCGGAACTACTCCGATCTCTCT 362
QY 175 uAlaTyTrpGlnGlnTyTrpAsnProGlnValGlyTyTrpCysArgAspLeuSerHisIleAla 195
DB 363 GGCAATATGAGAGATATAACCCGAGATGGGCTACTGAGGAGCCTGAGCCCATCGCGC 422
QY 195 alaPheLeuLeuTyTrpLeuProGlnGlnValAspAlaPheTrpAlaLeuValGlnLeuLeuAl 215
DB 423 CTGTGCTCTCTTATCTCTCTGAGAGAGATGATCTGGGCACTGGTGACGCTGCTGGC 482
QY 215 aSerGlnArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLe 235
DB 483 CAGTGAGAGGACATCTCTCGAGGAGATTTCAAGCCCAATATGGCGGACCGTCCAGGGGCT 542
QY 235 uGlnAspGlnGlnGlnIleValValAlaThrSerGlnProLysThrMetGlyHisGlnAs 255
DB 543 CCAAGACCAACAGAGCATGTGTAGCCACATCCCAACCCCAAGCATGTGGGGCATCGAGA 602
QY 255 pLysLysAspLeuLysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAs 275
DB 603 CAAGAAAGATTAATGTGGGCAGTGTCTCCGTTAGGCTGCTCATCCGAAATGTGATTGA 662
QY 275 pGlyIleSerLeuGlnGlyLeuThrLeuArgLeuTrpAspValTyTrpLeuValGlnGlyGln 295
DB 663 CGGATCTCTCTCGGGGCTACCCCTGGCGGCTGTGGAGCGTGAATCGTAGAAGAGCGAACA 722
QY 295 nAlaLeuMetPro-IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTr 315
DB 723 GCGGTAGATGCGCAATTACAAAGATCGCTTTAAGTTACACAGAAAGGCGCTCCAGAGA 782
QY 315 hr-SerArgCysGlyProTrpAla--ArgPheCysAsnArgPheValAspThrTrp-Ala 333
DB 783 GGTTCACGGTGTGCGCGTGGGGCCACGTTTTCACACGGGTTGTAATACCTGGGGCC 842
QY 334 ArgAsp-GlnAspThrValIleuLysHisLeuArgAlaSerMetLysLysLeuThrArgLys 353
DB 843 AGGCATTGAGGACACTGTGCTCAAGCATCTTAAGGCGCTCATGAAGAAATTAACAGAAA 902
QY 353 gGlnGly 355
DB 903 GCACGGG 909
RESULT 6
CD653466
LOCUS
DEFINITION AGENECOURT_14539069 NIA Human HI Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30423486 5', mRNA sequence.
ACCESSION CD653466
VERSION CD653466.1 GI:31891804
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Irene Glanis and Mahendra Rao, NIA
CDNA Library Preparation: Yulan Piao and Minoru Ko
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC
One distribution information
can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: NDAMS04 row: 5 column: 07
High quality sequence stop: 695.
Location/Qualifier

FEATURES

source

1. 868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30423486"
/issue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell CDNA Library (long)"
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTR, TRST, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and BOWS. When confluent (18-10 days
after plating), the ES cells from 4 x 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with TRIzol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID: 11544199] Double-stranded cDNAs were
synthesized with an oligo(dT) primer [Invitrogen:
5'-pGACTGATCTTCAGATCCGACGCGCCCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Loner-linker L1-SalI, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer SalI-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-Sport6 plasmid vector. The average insert size is
about 3.6kb."

ORIGIN

Alignment Scores:

Pred. No.: 1, 05e-86 Length: 868
Score: 1397.50 Matches: 267
Percent Similarity: 94.1% Conservative: 2
Best Local Similarity: 93.4% Mismatches: 15
Query Match: 46.8% Indels: 2
DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x CD653466 (1-868)

Qy 151 G|YThrLeuArglyVhHisIlePhePheArgAspArgTyG|YThrIysGlnArgGluLeu 170
Db 4 GGTAAATCCAAAGGACATATATCTTCAGGAGATCGATACGAAACCAAGCAGCGGAGAACTA 63

Qy 171 LeuHisIleLeuLeuAlaTyrgIuGlyTyAsnProGluValG|YTyCysArgAspLeu 190
Db 64 CTCACATCTCTCGGATATGAGAGATTAACCGAGGTGGCTACTCAGGAGACTG 123
Qy 191 SerHisIleAlaAlaLeuPheLeuLeuTyLeuProGluGluAspAlaPheTPrAlaLeu 210
Db 124 AGCCACATCGCGCGCTTGTTCCTCTCTATCTTCCTGAGGAGATGATTCCTGGGCACTG 163
Qy 211 ValGluLeuLeuAlaSerGluWatGHisSerLeuGlnGlyPheHisSerProAsnGly 230
Db 184 GTCCACCTGCTGGCCAGTGAGAGCACTCTCCAGGAGATTTCACAGCCCAATGGCGGG 243
Qy 231 ThrValGlnGlyLeuGlnAspGlnGlnGlnValValAlaPheSerGlnProIleThr 250
Db 244 ACCGTCAGGGGCTCCAGAGCCAAAGACATGTGTGACCGTACCAACCCAGAACACC 303
Qy 251 MetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIle 270
Db 304 ATGGGCGATCAGGACAAAGAAAGATCTATGTGGGCACTGTTCCCGTTAGGCTCCATC 363
Qy 271 ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTPrAspValTyLeu 290
Db 364 CGGATATTGATTGACGGGATCTCTCTCGGGCTCACCTCGCGCTGTGGACGTGTACTG 423
Qy 291 ValGluGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGln 310
Db 424 GTAGAAAGGCAAGCGCGTTGATGCGATTAACAAAGATGCTTTAAGTTACAGCAGAG 483
Qy 311 ArgLeuThrIleThrSerArgCysGlyProTPrAlaArgPheCysAsnArgPheValAsp 330
Db 484 CGGCTACAGAAAGCGTCCAGGTGTGGCCGCTGGGACGTTTTCACACCGGTCGTTGAT 543
Qy 331 ThrTPrAlaArgAspGluAspThrValLeuVhisLeuArgAlaSerMetLysValLeu 350
Db 544 ACCTGGCCAGGAGATGAGACACTGTGCTCAAGCATTTAAGGCGCTTAAGAAAGAACTA 603
Qy 351 ThrArgLysGlnGlyAspLeuProProProAlaLysProGlnGlnIleSerSerAlaSer 370
Db 604 ACAAGAAACAGGGGACCTGCACACCCCAAGCAACCCGAGCAAGAGTGTGGGATCC 663
Qy 371 ArgProValProAlaSerArgLysGlyLysThrLeuCysLysGlyAspArgGlnAlaPro 390
Db 664 AGGCGCTGTCGGCTTCACTGTGGCGGAGACCTTTCAGAGGGGAGCAGCGGCGCT 723
Qy 391 ProGlyProProAlaArgPheProArgProIleTPrSerAlaSer-ProProArgAlaPr 410
Db 724 CCAGGACCAACAGCCCGGTTCCCGGCAATGGTCAAGCTTCCCGCCACAGGGGACCC 783
Qy 410 ArgSerSerThrProCysProGlyGlyValAlaValArgGluAspThrTyPheValGln 430
Db 784 TCGTCTTTCACACCTGCTGTGGGGGCTGTC--NGSMAAACCTTACTGTGGGCTT 840
Qy 430 rGlnGlyValProSer 435
Db 841 CAGTGTGCCACCCGCC 856

RESULT 7
BUS42901 881 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10334768 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6574864
DEFINITION 5', mRNA sequence.
ACCESSION BUS42901
VERSION BUS42901.1 GI:22853384
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/BRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM2770 row: n column: 16
High quality sequence stop: 642.

FEATURES

Source

1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574864"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7, Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 7,486-83 Length: 881
Score: 1342.00 Matches: 267
Percent Similarity: 91.4% Conservative: 8
Best Local Similarity: 88.7% Mismatches: 17
Query Match: 45.0% Indels: 11
DB: Gaps: 2

US-10-071-838-2 (1-549) x BU542901 (1-881)

QY 81 MetLeuGlyAAspTgGluYrYrYsSerSerArGlyLeuIleAAspAglAATyrIlys 100
DB 1 TTGCTGGAGAGACTGGAGAGAA-TACAAAGCAGCAAGAAAGCTCATGATGAGACGTAAG 59
QY 101 GlyMetProMetAAsnIleArgGlyProMetTTPSerValLeuLeuAAsnIleGluGluMet 120
DB 60 GGAAATGCCCATGAACATCCGGGGCGGATGTGTGTCAGTCTCTGACACCTAGAGAAATG 119
QY 121 LysLeuLysAAspProGlyAArgTyrGlnIleMetLysGluYrYsArgSerSerGlu 140
DB 120 AAGTTGAAAACCCCGGAAGATACAGATCATGAAGAGAAAGAGTCACTTGAG 179
QY 141 HisIleGlnArgIleAAspAArgAAspValSerGlyThrLeuAArgLysHisIlePhePheArg 160
DB 180 CACATCCAGCGCATGACCGGAGACGTAACCGGAGACATTAAAGAGATATATCTTCAAG 239
QY 161 AAspAArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAATyrGluGluTyr 180
DB 240 GATTCATACGGAACCAAGCAGCGGGAATCACTCCATCTCTGCGCATATAGAGAGTAT 299
QY 181 AAspProGluValGlyTyrCysArgAAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
DB 300 AACCCGGAGGTGGCTACTGACAGGAGCTGAGCGACATGCGCGCTGTCTCTCTAT 359
QY 201 LeuProGluGluAAspAlaPheTrrAlaLeuValGlnLeuLeuAAspGluAArgHisSer 220
DB 360 CTTCTCGAGAGAGATTCATTTCTGGGCACTGTGTCAGCTCTGCGCACTAGAGGCACTCC 419
QY 221 LeuGlnGlyPheHisSerProAAsnGlyGlyThrValGlnGlyLeuGlnAAspGlnGlnGlu 240
DB 420 CTGCAAGGATTCACAGCCCAATGGCGGAGACGCTCAAGGGGCTCCAAAGCAACAGAGAG 479
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAAspLysLysAAspLeuCys 260

DB 480 CATGTGTAGCAACGTCAACAACCAAGACCATGGGCGATCAGAGCAAGAAAGATCTATGT 539
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAAspGlyIleSerLeuGly 280
DB 540 GGGGAGGTTCCTCCGTTAGGCTGTCTCATCCGATATTTGATGACGGGATCTCTCTGGG 599
QY 281 LeuThrLeuArgLeuTTPAAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB 600 CTCACCTTGGCGCTGTGGAGCGTATCTGTGTAGAGCGCAACAGGGGTTGATGCCATA 659
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArGlyPro 320
DB 660 ACAAGATTCGCTTTAAAGTTACAGCAGAGGCGCTCAGAAAGCTCAGAGTGTGGCCCG 719
QY 321 TTPAlaAArgPheCysAAsnAArgPheValAAspThrTTP-AlaAArgAAspLysAAspValLe 340
DB 720 TGGGCAAGTTTTCGCAACCGGTTCTGTGATACCTGGGGCGAGGATGAAGACACTGGTGC 779
QY 340 ULysHis---LeuAArgLAspMetLysLysLeuThrAArgLysGlnLysAAspLeuProPr 359
DB 780 TCAAGCAATCTTAAGGCGCTCTCTATGAAGAAAGCTTAACCAAGAAAGGCGGGGAGCC 839
QY 359 OProAlaLysProGluGlnGlySerSerAlaSerAArgProValProAlaSerArGlyGly 379
DB 840 CTG-GCAAAACCCAG-----CCACAACCCCGAAACCAAGAGCGG 877
QY 379 Y 379
DB 878 T 878

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1013)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM12766 row: 1 column: 15
High quality sequence start: 36
High quality sequence stop: 745.
Location/Qualifiers
1..1013
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744726"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.58e-82	Length:	1013
Score:	1338.50	Matches:	257
Percent Similarity:	93.4%	Conservative:	14
Best Local Similarity:	88.6%	Mismatches:	15
Query Match:	44.9%	Indels:	4
DB:	3	Gaps:	1

US-10-071-838-2 (1-549) x BMS60320 (1-1013)

```

Qy 1 MetApValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 98 ATGACCGTGTAGAGAGGTCTGTGTAGTTGTGTGGCACAAGGCCAGAGCATCTATTATG 157
Qy 21 LysTrpGluLysGlyValArgAlaGlyLeuProGluLysGlyProLysProPheArg 40
Db 158 AATATACAAAAGGACACCGAGCTGGCTGCCAGAGGACAAAGGGCTTAAGCTTTTGA 217
Qy 41 SerTrpAsnAsnValAspHisLeuGlyIleValHisGluTrpGluLeuProProLeu 60
Db 218 AGCTACAAACAACAATCATGATCTTTGGATGACATGAGACGAGCTGCTCTCTG 277
Qy 61 ThrAlaArgGluAlaValGlyLeuArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 278 ACTGCGGGGAGGAGGAAACAAATTCGGCGGAGATCAGCCAAAGAGCAATGGGTGAAA 337
Qy 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTrpLys 100
Db 338 ATGCTGGAGAGATGGGACACCTACAAAACAGCAAAAGCTCATAGATCGAGCTACAG 397
Qy 101 G1MetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGluMet 120
Db 398 GGAATTCCTCAATGAACATCCGGGGCCGATGTGGTCAGTCTCTGAACTTTGAGGAAATC 457
Qy 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 458 AAGTGAATAAAACCCCGAAGATACCATCATATGAAGAGAAAGGCAAGAGTCTATGAA 517
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 518 CACATCCAGAGATGAGACCTGACGTAAAGTGGAGCATTTAAGAGAGCATATATCTTCAGG 577
Qy 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
Db 578 GATCGATACGGAACCAAGACGGGAACTCTTTACATCTCCCTGGGGGTATGAGAGTAT 637
Qy 181 AsnProGluValGlyTrpCysArgAspLeuSerHisIleAlaIleAlaLeuPheLeuLeuTrp 200
Db 638 AACCCGAGAGTGGGCTACTGACAGGACCTGAGCCATCTCCGCTTCTCTCTTAT 697
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 698 CTTCCTGAGAGGATGATCTGTGGGCACTGGTGACGTGCTGCGCAATGAGAGGCACTCC 757
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 758 CTGAGAGGATTTCAACCCCAAAATGGCGGACCTTCAAGGGGCTCAAGACCAACAGAG 817
Qy 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAspLysLysAspLeuCys 260
Db 818 CATGTGTATGCCACGTACTACCAACCAATGTGGATCATGAGCAAGAAAGATCTATGT 877
Qy 261 -GlyGlnCysSer-ProLeuGlyCysLeuIleArg-IleLeuIleAsp--GlyIleSer 278
Db 878 GGGGCAAGTGTTCGCTCCTTAGGGCTGCTCATCCGGGATATTGATTGAATGGAGATCTCTCC 937

```

Qy 279 LeuGlyLeuThrLeuArgLeuTrp 286
 Db 938 TCGGGCTTCACCCCTTCGCTCTGG 961

RESULT 9
 BMS22600
 LOCUS
 DEFINITION BMS22600 919 bp mRNA linear EST 29-AUG-2001
 603175612P1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240090 5',
 mRNA sequence.
 ACCESSION BMS22600
 VERSION BMS22600.1 GI:15347392
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 919)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strauberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LHAM1605 row: 9 column: 03
 High quality sequence stop: 834.
 Location/Qualifiers

FEATURES

1. 919
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5240090"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.42e-82	Length:	919
Score:	1335.00 <td>Matches:</td> <td>247</td>	Matches:	247
Percent Similarity:	98.0% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	98.0% <td>Mismatches:</td> <td>4</td>	Mismatches:	4
Query Match:	44.8% <td>Indels:</td> <td>2</td>	Indels:	2
DB:	3	Gaps:	0

US-10-071-838-2 (1-549) x BMS22600 (1-919)

```

Qy 299 ProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTrpSerArgCys 318
Db 1 CCGATTAACAAGATCCCTTTAAGTTTCAGCAAGAGCGCTCAAGAGAGTCCAGGTGT 60
Qy 319 GlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThr 338
Db 61 GGGCCGGGGACAGTTTGTCAACCGGTTCTGTGATATCTGGGCGAGGTATGAGACACT 120
Qy 339 ValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuPro 358
Db 121 GTGCTACAGATCTTAGGGGCTCTATATGAATAAACAAGAAAGAGGAGGACCTGCAA 180

```


Qy	359	ProProAlaIalvProGIuNGIySerTerAlaSerArgProValProAlaSerArgIy	378
Db	181	CCCCAGCAACCCGAGCAAGGGTCGTGGCATCCAGGCTGTGGCGCTTCACGTGAC	240
Qy	379	GIyIvSThIuCyLbVgIyASpARgIInaIaProProGIyProProAlaIaArPhePro	398
Db	241	GGGAAGACCTCTCTCGAAGGGGACAGGAGGCCCTCCAGGCCACAGGCCGGGTTCCG	300
Qy	399	ArgProIleTPSerAlaSerProProArgAlaAPhArgSerSerThProCyAPProGIy	418
Db	301	CGGCCCATTTGATGACTTCCCGGCACGGGCACCTCGTTCTTCACACCTGTCCGT	360
Qy	419	GIyIaVaIaArgGIuASpThThTyProValGIyThrNGIyValProSerProAlaIeu	438
Db	361	GGGGTGTCGGAGAGACACTTACCTTGAGCACTAGGGGTGTGCCAGCCCGGCTTG	420
Qy	439	AlaGINGIyGIyProGINGIySerTPraRgPheIeuGINTPaINSerMetProARgIeu	458
Db	421	GCTCAGGAGAGACCTCAGGGTTCTCGAGATTCCTCGAGTGGAACTCATGCCCGCTC	480
Qy	459	ProThraPleuASpValIGIuGIyProTPPheArgHsIyIaSpPheArgINserCys	478
Db	481	CCAACGGACCTCGACGTAGAGGGCCCTTGTTCCGCCATYATGATTCAGACAGCGTCG	540
Qy	479	ThraValaRgaIaIeSerGINIuASpGINIuASpGINIuAProCYSTPGINAlaGIuHsIPro	498
Db	541	TGGTTCGTGCGCATATCCAGAGAGACAGACTGGCCCTCTGTCGAGGCTGAACACCTT	600
Qy	499	AlaGIuARgaIaArgSerAlaPheAlaIaProSerThraSPSerASpGINIyThr-Pr	518
Db	601	GCGAGCGGGGTGAGATTCGGCTTTGCTGTCCACCCAGCACTGATTCGACACAGGACCCCC	660
Qy	518	oPheARgaIaARgAPGIuGINProProCYAlaAPhThreXtIyProCYIleuCyGIGIyle	538
Db	661	CTTCAGACTTGGGAGCGAAGCCCTGTGTCTCCACCTCAGAGGCTTGCCTCGTGGCTC	720
Qy	538	uHsIleuGIuSerSerGINPheProProGIyPhe	549
Db	721	-CACTTGAAGATTCTCAGTTCCCTCCACGCTTC	753
RESULT 10			
LOCUS	B0920917		
DEFINITION	AGENCOURT_8926259 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462826		
ACCESSION	B0920917		
VERSION	B0920917.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini;		
TITLE	Homidae; Homo.		
JOURNAL	1 (bases 1 to 930)		
COMMENT	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LILCM2653 row: f column: 11 High quality sequence stop: 658. Location/Qualifiers 1..930 /organism="Homo sapiens" /mol_type="mRNA"		
FEATURES			
Source			

```

/db_xref="taxon:9606"
/clone="IMAGE:6462826"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH_MGC_101"
/note="Organ: Lung; Vector: pOTB7, Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGACG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

```

	Pred. No.:	1,429-81	Length:	930
Score:	1,324.00	Matches:	275	
Percent Similarity:	90.3%	Conservative:	5	
Best Local Similarity:	88.7%	Mismatches:	17	
Query Match:	44.4%	Indels:	17	
DB:	5	Gaps:	5	
US-10-071-838-2 (1-549) x BQ920917 (1-930)				
QY	162	ArgTgYGLYThrLySGlnAgtLnuLeuLnuHsIleLeuLeuAlaTyrgLnuLTrAAsn	181	
DB	3	CGATACGGAAACCG-CAGCGGAACTACTCCACATCTCTCCGGA-TATGAGAAAT-ACC	59	
QY	182	ProGLnuValGLYTYrCYsArpABLeuSeRnIsIleAlaIleuPheLeuTYrLeu	201	
DB	60	CCGAGAGTGGGCTACTGCGAGGACCTGAGCCACATCGCGCTTGTCCTCTATCTT	119	
QY	202	ProGLnuLnuApAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGLnuAgtHisSerLeu	221	
DB	120	CCTGAGAGGAGATGA-TTCTGGGCACCTGGGACGCTGGCCAGTGAAGAGGACCTCCCTG	178	
QY	222	GLnGLYpHeHisSerProAAsnGLYGLYThrValGlnGLYLeuGlnAAspGLnGLnuHis	241	
DB	179	CAGGATTTTCACAGCCCAATATGCGGAGCCGTCCAGGGGGCTCCAAGACCAACAGAGCAT	238	
QY	242	ValValAlaThrSerGlnProLYrThrMetGLYHsGlnAAspLYrLYsABLeuCYsGLY	261	
DB	239	GTGGTAGCCACGTCAACAAACCAAGACCAATGGGGCATCGAGCAAGAAAGATCTATGTGGG	298	
QY	262	GLnCYsSerProLeuGLYCYsLeuIleAgtIleuIleAAspGLYIleSerLeuGLYLeu	281	
DB	299	CAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTCACGGATCTCTCTGGGGCTC	358	
QY	282	ThrLeuAgtLeuTrpAAspValTYrLeuValGlnGLYGLnAlaLeuMetProIleThr	301	
DB	359	ACCTGCGCTGTGGGACGTGATCTGTAGAGAGCGAAAGCGGCTGTGATCCCATTAACA	418	
QY	302	ArgIleAlaPheLYrValGlnGLnuLYsAgtLeuTrpLYrThrSerArgCYsGLYProTrp	321	
DB	419	AGATTCCTCTTAAAGTTCAAGCAAGAGCCCTTCACAGAAAGAGCTCCAGGTGGCCCTGG	478	
QY	322	AlaAgtPheCYsAAsnAgtPheValAspThrTrpAlaArgAAspGLnuAspThrValLeuLYs	341	
DB	479	GCAGGTTTTTGCAACCGGTTGCTGTTGATCTCGGGCAGAGGATGAGAGCACGTGCTCAAG	538	
QY	342	HisLeuAgtAlaSerMetLYrLYsLeuThrAgtLYrGlnGLYAspLeuProProProAla	361	
DB	539	CATTTTAAAGGCTCTTATGAAGAACTTAACAAAGAAAGAGGGGAGCTTCACACCCCAAGCC	599	
QY	362	LYrProGLnuGLnuLYsSerAlaSerArgProValProAlaSerArgGLYLYrLYsThr	381	
DB	599	AAACCCAGCAAGAGGTCGTGGATTCAGAGCTGTGGCCGGCTTCACGTGGGGAGAACCC	658	
QY	382	LeuCYrLYsGLYAspAgtGlnAlaProProGLYrProProAlaAgtPheProAgtProIle	401	
DB	659	CTCTGCAAGGGGACAGGACGGCCCTTCAGAGGCCACACAGCCCGGTTCCCGCGGCCCATTT	718	

QY 402 TtpSer1AserProArgAlaProArgSerThrProCys-ProGlyGlyAla-V 421
 Db 719 TGGTCACTTCCCGCCACGGGACCTGTTCTTCCACACCTCTGTTGGGCTG 778
 QY 421 a1atrg1uapthTyrProValGly---Thrg1ng1yvalProSerProAlaLeu1ag 440
 Db 779 TCCGGGAAGACACCTACCTGTGTGGGAACCTTCAGGGGTGTCCAGCCCC-----GGGC 832
 QY 440 lng1yglYProG1ng1ySer-----TtpaYgPhleuG1nTtpaSer----- 454
 Db 833 CTGGGANTCAGGAAGAACCTCAGAGGCTTCCCTGAAA---ATTCCCTGCAGGGGGAAC 889
 QY 455 -----MetProArgLeuProThrAsp 461
 Db 890 TCATTGCCCCCTCCCAACGGA 915

RESULT 11
 BM553146 1141 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT 6542478 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742850
 DEFINITION 5', mRNA sequence.
 ACCESSION BM553146 GI:18791621
 VERSION BM553146.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 1141)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsbds-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM12761 row: k column: 11
 High quality sequence stop: 665.
 Location/Qualifiers
 1..1141
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5742850"
 /issue_type="medulla"
 /lab_host="DH10B"
 /clone_1b="NIH_MGC_119"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: ScaRI (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (ScaRI site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC library."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,21e-81 Length: 1141
 Score: 1333.00 Matches: 276
 Percent Similarity: 70.5% Conservative: 8
 Best Local Similarity: 68.5% Mismatches: 51
 Query Match: 44.4% Indels: 68
 DB: 3 Gaps: 6

US-10-071-838-2 (1-549) x BM553146 (1-1141)

QY 1 MetAspValValG1Val1aG1ySerTrpTrpAlaG1ng1uarg1uap1le1emc 20
 Db 102 ATGACGTTGTAGAGGTGCGGGCATGTGTGGGCACAAGACGACGACATCATTAATG 161
 QY 21 lyeTg1u1ySg1ySh1aGp1aG1yLeuProG1uap1ySg1yProPheArg 40
 Db 162 AANATCAAAAGGAGACCCAGCTGGCTGCCAGAGACAAAGGGCCTTAAGCTTTTTCGA 221
 QY 41 SerTyaanbanValaPhe1leuG1y1leValH1sg1uThrg1uLeuProLeu 60
 Db 222 AGTACAAACAACGTCGATTCATTGGGATTGTATCATGACGAGGCTGCTCTCTG 281
 QY 61 ThrAlaarg1u1a1ySg1n1learg1u1leSerArg1ySer1yTTPValaAsp 80
 Db 282 ACTGCGGGAGGGAAGCAAAATTCGGCGGAGATCGCCGAAGCAACAGTGGTGGAT 341
 QY 81 MetLeuG1yAspTrpG1u1yS1yT1yS1ySerSerArg1yS1e1aPArg1aT1yT1yS 100
 Db 342 ATCTGGGAGACTGGGAGAAATTCAAAAGACGAAAGCTCATAGATCGAGCTTACAG 401
 QY 101 GlyMetProMetAsn1learg1yPrometTrpSerVal1leu1euan1leG1u1umc 120
 Db 402 GGAATGCCATGAACATCCGGGCGCGATGTGTCACTCTCGGAACATGAGGAATG 461
 QY 121 lye1euan1yAsnProG1yArg1yG1n1leMet1ySg1ySg1y1yS1aTgSerSerG1u 140
 Db 462 AAGATGAAAAACCCCGGAATATACAGATCATGAAGAGGAAGGCAAGATCATCTGAG 521
 QY 141 H1e1leG1nArg1leAspArgAspValSerG1yThr1leuArg1ySh1lePhePheArg 160
 Db 522 CACATTCAGGCGCATTCACCGGACGTAAGCGGACATTAAAGACATTAATTTCTTACG 581
 QY 161 AspArgT1yG1yThrySg1nArg1u1euan1Sh1le1euan1aT1yG1u1uTyr 180
 Db 582 GATCGATACGGAACCAAGACGCGGAACTACTCTCAATCTCTCGCATATGAGGATAT 641
 QY 181 AsnProG1u1a1y1yT1yS1aGAp1e1SerH1e1a1a1euan1e1euan1e1yTyr 200
 Db 642 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCAACATCCCGCTTGTCTCTCTAT 701
 QY 201 LeuProG1u1uap1aPheTrpAlaLeuValG1leu1euan1a1aSerG1u-ArgH1se 220
 Db 702 CTTCCTGAGGAGATGATTCATTCGGGACCTGGTCAAGTCTGTCAGTGAAGGCAATTC 761
 QY 220 r1euan1y1yPheH1se1yProan1yG1yThrValG1ng1yLeuG1nAspG1ng1 240
 Db 762 CTGCAAGGATTTCAAGCCCAATGGGGGACCGCCCGGGGCTCCAGAACCCACAGCA 821
 QY 240 wh1sVal1a1aThSerG1nPro1yThrMetG1yH1eG1nAsp1y1yS1aP1e1uCy 260
 Db 822 GCATGGGGGAACCCGCTTACACCCGACCCCTGGG----- 858
 QY 260 sG1yG1nCySerProLeuG1yCyS1euan1earg1leuan1easG1y1leSerLeuG1 280
 Db 858 ----- 858
 QY 280 yLeuThrLeuArgLeuTrpAspVal1y1r1euan1a1G1uG1u1a1euan1euan1e 300
 Db 858 ----- 858
 QY 300 ethrArg1lealAph1e1yVal1G1n1ySarg1e1uThr1yTThSerArgCySg1yPr 320
 Db 859 -----GACTCAAAAGCCCTCACCAAGATCCCAATTTGTTGGCC 896
 QY 320 oTTPAlaArg---PheCyAsnArgPheValAspThr-TTPAlaArgAspG1uap1yThr- 338
 Db 897 CCGGCGCCCTTCTTACAAAGCGGTTCTTTGTACTTGGGTCAGGCAATCGGGGACCC 956
 QY 339 --Val1euan1ySh1a1euan1a1aSerMet---Lye1yS1euan1yThrg1ySg1n1yAsp1 357
 Db 957 CGGTGCTCAAGACCTTTTGGGCTCTCTGTAATAAACTACCCGGAAGAAAGCAAGGGGG 1016

Oy	369	1aseArgPrvValProAlaSerAglGlyLysThrLeucylAsparGlna	389
Db	1077	CCTGGCACCCCTTTTCCCGCGCGCGGAGAAATCACCACCGCGCGGGAACAAGTA	1133
Oy	389	1apRo 390	
Db	1137	TCCCA 1141	
RESULT 12			
BMS63571			
LOCUS			
DEFINITION	BMS63571	1079 bp	mRNA linear EST 20-FEB-2007
ACCESSION	AGENCOURT_6564823 NIH_MGC_119 Homo sapiens	cdna clone IMAGE:5744077	
VERSION	5', mRNA sequence.		
KEYWORDS	BMS63571.1 GI:18810609		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1079)		
AUTHORS	NIH-MGC http://mhc.mci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: LLML2764 row: n column: 14 High quality sequence step: 665.		
FEATURES			
source	location/Qualifiers		
	1..1079		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5744077"		
	/tissue_type="medulla"		
	/lab_host="DH10B"		
	/clone_1lb="NIH_MGC_119"		
	/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note this is a NIH_MGC Library."		
ORIGIN			
Alignment Scores:			
Pred. No.:	5..38e-80	Length:	1079
Score:	1302.50	Matches:	262
Percent Similarity:	91.9%	Conservative:	15
Best Local Similarity:	88.8%	Mismatches:	9
Query Match:	43.7%	Indels:	1
DB:	3	Gaps:	
US-10-071-838-2 (1-549) x BMS63571 (1-1079)			
Oy	1	MetAepValValGIUValAlagIySeRTPTPALgInGUArGUluePtlletMet	20
Db	100	ATGGAAGTGTCGTGAGGTCGCCGGGTATTGTGGGCAAGAAGGAGAACATCATATTATG	159

OY		21	LysrTgcluysgllyhi saargal aglyleuProglubaprysglyProlysPPropeAarg	40
Db		160	AAATACGAAAAAGGACACCAGACTGGGCTCCAGAGGACAAGGGCGCTTAAGCCTTTTCGA	219
OY		41	SerTyrsAmnbnAnvAlaAPHtIsleUglyIleValHisgluthrgIuleuProProLeu	60
Db		220	AGCTACCAACAAACGTCATCATTTGGGGATTGTACATGACCGAGCTGCCCTCTG	279
OY		61	ThrAlaarglualalysglnlleatgarlglnlleserTgylsserySTPvalAsp	80
Db		280	ACTCGCGGGAGGGCGAAGCAAATTCGCGGGGAGTCAAGCCGAAGAACAAGTGGTGAGAT	339
OY		81	MetIeuGlYAAPTTPgilylyrTyLysSerSerglyysleuIleaAPalaTYLYs	100
Db		340	ATGCTGGAGACTGGGAGAAATACAAAAGCACAGAAAGCTCATNGATCCAAGCTACAG	399
OY		101	GlyMetProMeCaenilleargglyPrometTrpSerValIeueuanilleglumet	120
Db		400	GGATGGCCCATGAAACATCCGGGGCCGATGGTCACTCTCTGAACATTGAGGAATG	459
OY		121	LytleuLyAsnP-rogiyArtyrGlnIlleMelLyglulysglylysaagsSerSglu	140
Db		460	AAGTTGAAAAACCCCGAAGATACACAGATCATGAAAGAGAGGACAAGAGTCACTTCAG	519
OY		141	HistIlegIntrgIleAParAPesArgyleuHisIleleuEualaTyrgluTyrr	180
Db		520	CACATTCAGCGCATCGACCGGGACATAAGCGGCATTTAAGAAACATATGTTCTTAG	579
OY		161	AsparTgryrglyThrylysglnargIuleuIleHisIleleuEualaTyrgluTyrr	180
Db		580	GATCGATACGGAACCAAGACAGCGGAGACTCTCACACTCTCGGCATATGAGAGAT	639
OY		181	AsnProGIuValIGlyTyrcysArgAspLeuSer-HisIIlaIAlaleuheuHeuLeuTy	200
Db		640	AAACCGAGGGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTTCTCTCTTA	699
OY		200	rLeuProGIuLIuAspAlaPhetPALaleuValIGlnIleuIleuIleasergIuarghiase	220
Db		700	TTTTCTCGAGAGAGATCATTTCTGGGACCTGTGTCACTGCTGSCAGTGAGAGGACTC	759
OY		220	rLeuglnglyPhehisserProaen-GlyglyThrvAlGlnglyLeuGlnAbpIingng	240
Db		760	CCTGACAGGATTTTCAAGCCCAAAATGGCGGGAGCTCAGGGGGCTCCAGACCAACAG	819
OY		240	Iu-HisValValAla---ThrSeGIuProlysrHrmetGlyHisgin-AsplyslyAs	258
Db		820	AAGCATGGGGTAGCACACGTTCAACATCAAGACCAATGGGGGCATTCAGGACAGAAAGA	879
OY		258	P-LeucYsglyGlnCyseSerPro-IeuglyCyaleuileargIleleu--IleApqlyI	277
Db		880	TCTCATGtGGGCAAGTGTTCGCGGTAGGGGTGCCCTATCCGATAAATTGATGAGCGGGA	939
OY		277	IeSerLeuGlyLeuThrIeuarIgleUTP	286
Db		940	TCTCTCTTCGGGTTCAACCCCTGCGCGCTGG	968
<hr/>				
RESULT 13				
B1828272		831 bp	mRNA	linear EST 04-OCT-2001
LOCUS		603078163p1	NIH_MGC_119 Homo sapiens cDNA clone IMAGE:516999 5'	
DEFINITION		mRNA sequence.		
ACCESSION		B1828272		
VERSION		B1828272.1	GI:15939822	
KEYWORDS		EST.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
		Homidae; Homo.		
REFERENCE		NIH-MGC http://mgc.nci.nih.gov/		
AUTHORS		1 (bases 1 to 831)		
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)		

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1422 row: n column: 16
High quality sequence stop: 822.
Location/Qualifiers

FEATURES

source

```
1..831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5169999"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

Score:	4.5e-78	Length:	831
Percent Similarity:	1272.50	Matches:	261
Best Local Similarity:	93.0%	Conservative:	3
Query Match:	91.9%	Mismatches:	10
	42.7%	Gaps:	16
			2

US-10-071-838-2 (1-549) x B1828272 (1-831)

QY 148 AspValSerGlyThrLeuArgLysHisIlePhePheArgPheArgTyrGlyThrLysGln 167
 Db 2 GACATAGCGGACCATTAAGAAAGCATATGTTCTTCAGGATGATACGAAACCAACAG 61
 QY 168 ArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyrAspProGluValGlyTyrCys 187
 Db 62 CGGGAACACTCCACATCTCTCTGGCATATGAGAGATATACCGGAGGTGGGCTACTGC 121
 QY 188 ArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPhe 207
 Db 122 AGGAGACTGAGCACAATCGCCGCTTGTCTCTCATCTTCCTCGAGGATGATTC 181
 QY 208 TrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerPro 227
 Db 182 TGGGACCTGCTGAGCTGCTGGCAGTGAAGGACACTCCTGCAAGGATTTCAAGCCCA 241
 QY 228 AsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGluHisValValAlaThrSerGln 247
 Db 242 AATGGCGGACCGTCCAGGGGCTCCAAAGCAACAGAGCATGTGGTACGCTACCA 301
 QY 248 ProLysThrMetCysLysGlnAspLysGlyGlnCysSerProLeuGly 267
 Db 302 CCCAACAACATGGGGATCAGCAAGAAAGATCTATGTGGGAGTGTCCCGTTAAGGC 361
 QY 268 CysLeuLeuArgIleLeuLeuAspGlyLysSerLeuGlyLeuThrLeuArgLeuTrpAsp 287
 Db 362 TGCTCATCCGATATTTGATGACGGGATTTCTCTGGGCTTCACTCCGCTGTGGAC 421
 QY 288 ValTyrLeuValGluGluGlnAlaLeuMetProIleThrArgIleAlaPheLysVal 307
 Db 422 GTGTATCTGGTAGAAGCGAAAGCGGCTTGAATCCGATTAACAAGAAATCCGCTTAAAGTT 481

QY 308 GlnGluLysArgLeuThrLysSerArgCysGlyProTrpAlaArgPheCysAsnArg 327
 Db 482 CAGCAAGAGCGCTCCAGAAAGACGTCAGAGTGTGGCCGTGTGGACATTTTTCACACCGG 541
 QY 328 PheValAspThrTrpAla-ArgAspGluAspThrValIleLysHisLeuArgAlaSerMe 347
 Db 542 TTGCTTATACCTGGGCGCCAGGATGAGACACTGCTCAAGCATTTTGGGCTCTTAT 601
 QY 347 LysLysLeuThrArgLysGlnGlyAspLeuProProProProAlaLysProGluGlnGly 367
 Db 602 GAAGAAACTAACAAGAAAGCAGGCGGACCTGCCACC-CCAGCCAAACCCGAGCAAGGCTC 660
 QY 367 rSerAlaSerArgProValProAlaSerArgLysGlyLysThrLeuLysGlyAspArg 387
 Db 661 GTGGCATTCAGGCTGTGTGCGCTTCACGTGGGGAAGACCTCTGCAAGGGGGAGCA 720
 QY 387 GlnAlaProProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProPr 407
 Db 721 GCGAGGCTTCAGGCCCAAC-GCCCGTT-CCGCGGCATT-TGATCAGCTTC-CCGCC 776
 QY 407 GATGAlaProArgSerSerThrPro-----CysProGlyGlyAlaValArgG1 423
 Db 777 ACGGGC-CTCGTTCTTCCAAACCTGTCTGTGGGCTGTCCG-----GA 820
 QY 423 uAspThrTyr 426
 Db 821 AGACACCTTAC 830

RESULT 14

BI088323

LOCUS

DEFINITION

602852793F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4994398 5',

mRNA sequence.

ACCESSION

BI088323

VERSION

BI088323.1 GI:14506653

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 709)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LHAM1016 row: 1 column: 23

High quality sequence stop: 709.

Location/Qualifiers

1..709

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4994398"

/cell_line="MGC36"

/lab_host="DH10B"

/clone_lib="NIH_MGC_10"

/note="Organ: Cervix; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.5 kb. Library prepared by Life

Technologies.."

ORIGIN
Alignment Scores:

Pred. No.: 3.99e-76 Length: 709
 Score: 1243.00 Matches: 235
 Percent Similarity: 99.6% Conservative: 0
 Best Local Similarity: 99.6% Mismatches: 1
 Query Match: 41.7% Indels: 0
 Gaps: 2

US-10-071-838-2 (1-549) x B1088323 (1-709)

QY 47 AapHleUGLYlIleValHleGluThrgluLeuPProleuThralAargluAlay 66
 DB 2 GATCATTTGGGGATTGTACATAGACGGAGCTGCTCTGACACGGCGGAGCGGAG 61
 QY 67 GlnleAargGluIlleSerArglySerlySTPValAapMetleuGlyAapTrglu 86
 DB 62 CAAATTCGGCGGAGATCAGCCGAAAGACAAAGTGAGATGCTGGAGACTGGAG 121
 QY 87 LysTrlySerSerArglySerlyleAapArgAlaTylyGlyMePProMetAanle 106
 DB 122 AAATTCMAAAGACAGAAAGCTCATAGATCGAGCTCAAGGAATGCCATGAACATC 181
 QY 107 ArgGlyPProMetTrpSerValleuLeuAnilleGluGluMetlyleuLysAnPProGly 126
 DB 182 CGGGGCGCGATGTGGTCACTCTCTGAACTGAGAAATGAAATGAAAAACCCCGGA 241
 QY 127 ArgTrgGlnlleMetlyleuGlylyArglySerSerGlnHleGlnAarglleAap 146
 DB 242 AGATACCAAGATCATAGAGAGAAAGAGAGATCATGACACATCCAGCGCATCGAC 301
 QY 147 ArgAapValSerGlyThrlleuArglylyHlePhePheArgAapArgTrglyThrllys 166
 DB 302 CGGACGTAAAGCGGACGCAATTAAGAAAGCTATATCTTCAGGATCGAAGCAAGCAAG 361
 QY 167 GlnArgGlyleuLeuHleleleuLeuAlaTylyGluGluTyAsnProGluValGlyTr 186
 DB 362 CAGCGGAACTACTCACAATCTCTCGCATATGAGATATTAACCCGAGGTGGGCTAC 421
 QY 187 CysAapAapLeuSerHleleleleuPheleuLeuTyTrleuProGluGluAapAla 206
 DB 422 TGCAGGAGCTTAGCCACATCCCGCTTGTCTCTCTATCTCTGAGAGAGATGCA 481
 QY 207 PheTrpAlaLeuValGlnleuLeuAlaSerGluArgHleSerleuGlnGlyPheHleSer 226
 DB 482 TTCTGGGCACTGTGTCAAGCTCTGGCCAGTGAAGCACTCTCGAGGAGATTTCACAGC 541
 QY 227 ProAanglyGlyThrlValGlnGlyleuGlnAapGlnGlnHleValAlaThrSer 246
 DB 542 CCAATGGCGGAGCGGTCCAGGGGCTCCAAAGACCAAGAGACATGTGTACCCACGTCA 601
 QY 247 GlnPProlyThrlMetGlyHleGlnAaplylyAapLeuCysGlyGlnCysSerProleu 266
 DB 602 CAACCCAAAGACATGGGGCATCAGACAAAGAAAGATCTATGTGGGAGTGTCCCGGTTA 661
 QY 267 GlyCysleuIleArglleuIleAapGlylleSerleuGlyleuThr 282
 DB 662 GGCTGCTCATCCGATATTGATGACGGGATCTCTCGGCTCACCC 709

RESULT 15
 BQ891586 946 bp mRNA linear EST 16-AUG-2002
 LOCUS AGENCOURT 8670678 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6580487
 DEFINITION 5' mRNA sequence.
 ACCESSION BQ891586
 VERSION BQ891586.1 GI:22283600
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 946)
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rgs@ds-rsmail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNC9569 row: c column: 16
 High quality sequence start: 2
 High quality sequence stop: 554.
 Location/Qualifiers
 1..946
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6580487"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOT87; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GCGACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH-MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 8.84e-76 Length: 946
 Score: 1240.50 Matches: 259
 Percent Similarity: 85.2% Conservative: 12
 Best Local Similarity: 81.4% Mismatches: 32
 Query Match: 41.6% Indels: 15
 Gaps: 5

US-10-071-838-2 (1-549) x BQ891586 (1-946)

QY 82 LeuGlyAapTrp-GluGlyTrlySerSerArglyleuIleAapArgAlaTylyGly 101
 DB 4 TTGGGAGACTGGGGAATATCAAAAGACAGCAAGAGCTCATGATCGAGCGTCAAGGG 63
 QY 101 YMePProMetAanilleAarglyPProMetTrpSerValleuLeuAnilleGluGluMetly 121
 DB 64 AATGCCATGAAACATCCCGGGCCGATGTGTCACTCTGAAACACTGAGAAATGAA 123
 QY 121 ALeuLysAnPProGlyArgTrgGlnlleMetlyGluGlylyAapSerSerGlnHle 141
 DB 124 GTTGAANAACCCCGGAATATCCAGATCATAGAGAAAGGCAAGAGTCTATGAGCA 183
 QY 141 AileGlnAarglleAapArgAapValSerGlyThrlleuArglylyHlePhePheArg 161
 DB 184 CATCAGGCAATTCAGCCGGAAGTGAAGGAGCAATTAAGAAAGCATATATCTTCAGAGGA 243
 QY 161 PArgTrglyThrllyGlnAargGluLeuLeuHleleleuLeuAlaTylyGluGluTyAs 181
 DB 244 TCGATACCGAACAAGACGCGGAATCTACTCCTCTGCGATGTGAGAGATATA 303
 QY 181 nPProGluValGlyTrlyCysAapAapLeuSerHleleleleuPheleuLeuTyTrle 201
 DB 304 CCGGAGGTGGCTACTCAGAGACCTAGACCATCCCGCTTGTCTCTCTACT 363
 QY 201 nPProGluLysAapAapTrpAlaLeuValGlnleuLeuAlaSerGluArgHleSerle 221
 DB 364 TCCTAGAGAGAGATGATCTCGGACCTGTGCAAGTCTGGCCAGTGAAGCACTCCCT 423
 QY 221 uGlnGlyPheHleSerProAanglyGlyThrlValGlnGlyleuGlnAapGlnGlnHle 241
 DB 424 GCAGGAGATTCAACGCCAAATGGCGGAGCGTCCAGGGGCTCCAAAGACAAAGAGACA 483

QY 241 sValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCysG1 261
|||
Db 484 TGTTGTTAGCCAGTCAACCAAGACCATGGGCGATCGAGACAGAAAGATCTATGTGG 543
QY 261 YGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLe 281
|||
Db 544 GCAGTGTGTCCTGGTACGCTCCATCCGATATTGATGACGGGATCTCTCGGGCT 603
QY 281 uThrLeuArgLeuThrAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIleTh 301
|||
Db 604 CACCTGCGGCTGTGGGACGTGTATCTGTAGAAAGCGCAACAGCGGTGATGGCATTAAC 663
QY 301 rArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgCysGlyPro-T 321
|||
Db 664 AAGATCGCTTTAAGGTTGACAGAAAGCGCTCAACCAAGACGGCGAGGGGGCCGGG 723
QY 321 rPalArg-PheCysAsnArg--PheValAspThrTrpAlaArgAsp---GluAspThr- 338
|||
Db 724 GGGCAGCTTTTTCACACCGGGGTCTTGATACCTGGGGCAGGGGATGAGAGCACTG 783
QY 339 --ValLeuLysHisLeu-ArgAlaSerMetLysLysLeuThrArg----- 352
|||
Db 784 GGGGTTAAAGCCATCTTAGGGGCTCTTATGAGAAACTAAACCAAGAAAGAGAGG 843
QY 353 ---LysGlnGlyAspLeuProProProAlaLysProGluGlnGlySerSerAlaSerArg 371
|||
Db 844 GGGGAACCTGGCAAAACCCCAAGCCCAAAACCCCAAGCAAGGGGGGTCTGGTTCCG--- 900
QY 372 ProValProAlaSerArgGlyGlyLysThrLeuCysLysGly 385
|||
Db 901 GCAATTCCCAAGGGCGGTGGGGGCGGGGCTTTCCCGCGGG 942

Search completed: April 5, 2006, 16:26:30
Job time : 6220 secs

This Page Blank (uspto)

PA (TULIA-) TULARIK INC.

XX L.J. J., Powers S., Xiang P., Peng Y.

XX WPI; 2002-706902/76.

DR P-PSDB; ABG70736.

XX Novel isolated PRCl7 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRCl7 activity
or expression relative to normal, for example cancer.

XX Claim 24; Page 62; 78pp; English.

CC The present invention relates to a new PRCl7 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRCl7 activity or expression relative to normal, for example
CC cancer. The present nucleic acid sequence represents the human PRCl7 gene
CC located on chromosome 17q11-12. This sequence encodes the human PRCl7
CC protein of the invention

XX Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.72e-128 Length: 1964
Score: 2983.00 Matches: 549
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-071-838-2 (1-549) x ABS54706 (1-1964)

QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 1 ATGACAGCTGTCAGAGCTCGCGGCACTTGTGGCACAAGCCAGAGGAGCATCTATATG 60
QY 21 LysTrpGluLysGlyHisAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 61 AAATAACAAAAGGAGACACCGAGCTGGCTGCCAGAGGACAAAGGGCCCTTAAGCCTTTTGA 120
QY 41 SerTrpTrpAsnAsnValAlaAspHisLeuGlyIleValHisGlyIleTrpGluLeuProPhe 60
DB 121 AGCTTACAAACAACAGCTGATCTTGGGATTGTACATGAGACGAGCTGCCCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 181 ACTCGCGGAGGCGGAGCAAAATTCGCGGAGATCACCCGAAAGAGCAAGTGGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 241 ATGCTGGGAGACTGGAGAAATTAACAAGCAGAGAAAGCTCATATGATCGAGCTTACAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTCACTCTTCGAACTTGAAGAAATG 360
QY 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 361 AAGTTGAAAACCCCGGAGAGTACCAAGTCAATGAAGAGAGGAGGAGGTCTATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 421 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAGAGACATATATCTTCAGG 480
QY 161 AspArgTrpGlyThrLysGlnArgGluLysLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 481 GATGATACGGAAACCAAGCAGCGGAGAACTACTCCACATCTCTGGGCAATAGAGAGAT 540

QY 181 AsnProGluValAlaGlyTyrCyAspAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
DB 541 AACCCGAGGTGGGCTACTCGAGGAGCTGAGGACATCGCCCTTGTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluAspHisSer 220
DB 601 CTTCCTGAGAGAGATGATCTGGGCACTGGTCACTGTGTGGCCAGTGAAGAGCTCC 660
QY 221 LeuGlnIlePheHisSerProAsnGlyIleThrValGlnIleLeuGlnAspGlnGlu 240
DB 661 GTGAGGAGATTTTCAACGCCAAATGGCGGAGCCGTCCAGGGGCTCCAAAGCCAAAGAG 720
QY 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAspLysLysAspLeuCys 260
DB 721 CAGTGGTATCCAGCTGACAAACCCAGACATGGGCGATCAGGACAAAGAAAGTATATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 781 GGGCAGGTTCCTCCGTTAGGCTCCATCCGATTTGATTTGACGGGATCTCTCGGG 840
QY 281 LeuThrLeuAspLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
DB 841 CTCACTCTGGCTGTGGAGCTGTATCTGTGAAAGGCAAGGCGTTGATCCGAT 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 901 ACAAGATCGCTTTAAGTTAGTTAGAGAGCCGCTCCAGAAAGCTCCAGGTGTGGCCG 960
QY 321 TrpAlaArgPheCysAsnAspPheValAspTrpTrpAlaArgAspGluAspTrpValLeu 340
DB 961 TGGCAGCTTTTGGCAACCGGTTCTGTGATACCTGGGCGGAGATGAGCACTGTGCTC 1020
QY 341 LysHisLeuAspArgLysSerMetLysLysLeuThrArgLysGlnIleAspLeuProPro 360
DB 1021 AAGCATTTTGGGCTCTTAAAGAACTAACAAGAAAGAGGAGGAGCTGCCACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1081 GCCAAACCCGAGCAAGGAGTGTGGCATCCAGGCTGTGGCTTCACTGGCGGAGAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1141 ACCCTCTGCAAGGAGGAGCAGAGGCGCTCCAGGCCCAACAGCCGCTTCCGCGGCC 1200
QY 401 IleTrpSerAlaSerProProAlaProArgAlaProArgSerSerThrProCysProGlyAla 420
DB 1201 ATTTGGTCACTTCCCGGACCGGACCTGTTCTTCCACACCTGTCTGTGGGGCT 1260
QY 421 ValArgGluAspTrpTrpProValGlyThrGlnIleValProSerProAlaLeuAlaGln 440
DB 1261 GTCGCGGAGAGACACTTACCTGTGGGACATCAGAGGTGTGCCAGCCGCGCTGGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProAlaGluProThr 460
DB 1321 GAGAGACTTCAAGGATTCCTGAGATTCCTGAGATGAACTCAATGCCCTCCCAACG 1380
QY 461 AspLeuAspValAlaGluIleProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1381 GACCTGACCTTAAGAGGCGCTTGTGCTTCCGATTAATTCAGACAGAGCTGTGGGT 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1441 CGTGCCATATCCAGAGGAGCAGAGCTGGCCCTGTGGAGGCTGAACACCTGCGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1501 CGGATGAGATCGGCTTGTGTCACCCAGCATGATTTCCACACGAGGACACCCCTTCAG 1560
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1561 GCTAAGGACCAACAGCGGTGTGCTCCACCTCAGGGGCTTGTGGGCTTCCACTTG 1620
QY 541 GluSerSerGlnPheProGlyPhe 549

Db 1621 GAAAGTTCTCAGTTCCTCCAGGCTTC 1647

RESULT 2

ADCC37382

ADCC37382

18-DEC-2003 (first entry)

Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 215.

Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;

neurodegenerative disease; bone disease; AIDS; Anti-inflammatory;

immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;

Neuroprotective; Nootropic; Cardiac; Gene therapy; human; gene; de-

Homo sapiens.

MO2003048202-A2.

12-JUN-2003.

03-DEC-2002; 2002MO-JP012644.

03-DEC-2001; 2001JP-00368692.

05-DEC-2001; 2001US-0358282P.

03-OCT-2002; 2002JP-00291302.

04-OCT-2002; 2002US-0415769P.

(ASAH) ASAH KASEI KK.

Matsuda A, Muramatsu S;

WPI; 2003-505282/47.

P-PSDB; ADCC37383.

New purified protein that activates nuclear factor kappa B (NF-kappaB),

useful for treating inflammation, autoimmune diseases, cancers,

infectious diseases, bone diseases, AIDS, neurodegenerative diseases or

ischemic disorders.

Claim 4; SEQ ID NO 215; 938bp; English.

The present invention relates to novel proteins and their coding

sequences (ADCC37168-ADCC37455), which activate nuclear factor kappa B (NF-

kappaB). The proteins and their coding sequences are useful for treating

a disease associated with NF-kappaB activation, such as inflammation,

autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,

neurodegenerative diseases, or ischemic disorders.

Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,72e-128	Length:	1964
Score:	2983.00	Matches:	549
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-071-838-2 (1-549) x ADCC37382 (1-1964)

Qy	1	MetAspValAlaGluValAlaGluSerTyrTrpAlaGlnGluArgGluAspIleIleMet	20
Db	1	ATGACCTGCTAAGAGTCCGCGGAGTGGGACAAAGCAGAGGACATCATTTATG	60
Qy	21	LYeTGTGluYsGluYsIleAspAlaGluLeuProGluAspIleYsGluYsProIleAsp	40
Db	61	AAATACGAAAGGAGACACGAGCTGGGCTGCCAGAGGACAAAGGGGCTTAAGCTTTTGA	120

Qy	41	SerTyrAsnAsnAsnValAspHisIleuGluYleValHisIleGluThrGluLeuProProIleu	60
Db	121	AGCTACAAACAACAGCTCGATCATTTGGGATTTGTACATAGACGAGCTGCTCTCTG	180
Qy	61	ThrAlaArgGluAlaYsGlnIleArgArgGluIleSerArgYsSerIleYsTrpValAsp	80
Db	181	ACTGCCGGAGGAGGAAAGAAATTCGCGGAGATACGCCGAAAGACAAAGTGGTGGAT	240
Qy	81	MetLeuGluYsAspTrpGluYsYrYsSerSerArgYsLeuIleAspAlaGluYs	100
Db	241	ATGCTGGAGACCTGGGAAATACAAAGAGCAGAAAGCTCATATTCAGCTACAG	300
Qy	101	GlyMetProMetAsnIleArgGluYsProMetTyrSerValLeuMetAsnIleGluMet	120
Db	301	GAAATGCCATGAAACATCCGAGGCCGATGTGCTCTCTGAAACATTTAGGAAATG	360
Qy	121	LYeLeuYsAsnProGluYsArgYsGlnIleMetYsGluYsGluYsArgYsSerGlu	140
Db	361	AAAGTGAAGAAACCCGGAATACCATGATCATGAGGAGAAAGGCAAGAGGTACTGAG	420
Qy	141	HisIleGluArgIleAspAlaGluValSerGluYsThrLeuArgYsHisIlePhePheArg	160
Db	421	CACATCCAGCGCATTCGACCGGAGCTTAAGCGGACATTAAGAGCATATATCTTCAGG	480
Qy	161	AspArgTyrGluYsThrYsGluArgGluLeuHisIleLeuMetAlaYsGluGluYs	180
Db	481	GATCGATACGAAACCAAGCGGAACTCTCCACATCTCTGGGATATAGAGATAT	540
Qy	181	AsnProGluValGluYsYsArgYsAspLeuSerHisIleAlaIleLeuPheLeuYs	200
Db	541	AACCCGAGGTGGGCTACTCGAGGACCTAGCCACATCCCGCTTCTCTCTAT	600
Qy	201	LeuProGluGluAspAlaPheTrpAlaLeuValGluLeuLeuAlaSerGluArgHis	220
Db	601	CTTCCTGAGGAGATGATCTGGGACCTGGTCAAGCTCTGGCCAGTGAAGGCACTCC	660
Qy	221	LeuGluGluYsPheHisSerProAsnGluYsIleValGluGluGluArgGluGlu	240
Db	661	CTCGAGGATTTTCAAGCCCAATGGGGGACGCTCCAGGGCTCCAAAGCAACAGAG	720
Qy	241	HisValValAlaThrSerGluProYsThrMetGluYsGlnAspYsYsAspLeuYs	260
Db	721	CATGGGTAGCAGCTGACCAACCAAGACCATGGGGCATGAGGACAAAGAAATCATGT	780
Qy	261	GlyGluYsSerProLeuGluYsLeuIleArgIleLeuIleAspGluYsSerLeuGlu	280
Db	781	GGGCAAGTGTCCCGTTAGGCTGCTCATCCGATTTGATTGACCGGATCTCTCGGG	840
Qy	281	LeuThrLeuArgLeuTrpAspValYrLeuValGluGluGluGluGluMetProIle	300
Db	841	CTACCCCTGGCTGTGGACGTGTATCTGTGAAGGCGAAACAGGCGTGTATGCGATA	900
Qy	301	ThrArgIleAlaPheYsValGluGluYsArgLeuThrYsThrSerArgYsGluYs	320
Db	901	ACAAGATGCCCTTTAAGGTTACGACGAAGCGCTCACAAAGCTCCAGGTGGCCCG	960
Qy	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal	340
Db	961	TGGGCAAGTTTGGCAACCGGCTCGTTGATACCTGGGCGAGGATGAGCACGTGCTC	1020
Qy	341	LYeHisLeuArgIleAspMetYsYsLeuThrArgYsGluGluYsLeuProProPro	360
Db	1021	AMGCATCTTAGGGCTCTTATGAAGAACTTACAAAGAAAGGAGGAGCTGCCACCCCA	1080
Qy	361	AlaYsProGluGluGluYsSerSerAlaSerArgProValProAlaSerArgGluYs	380
Db	1081	GCCAAACCGAGCAAGGCTGTGCGCATCAGGCTGTGCGGCTTCACTGCGGGAAG	1140
Qy	381	ThrLeuYsYsGluYsAspArgGlnAlaProProGluYsProProAlaArgPhePro	400
Db	1141	ACCCTCTGCAAGGAGGAGCAGGCGCCCTCCAGGCCCAAGCCCGGCTTCCCGGCCCC	1200
Qy	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGluYsAla	420

```
Db 1201 ATTGGTCAGCTCCCGCCACAGGACCTGGTCTTCACACCCCTGCTGAGGAGCT 1260
Qy 421 ValArgLysAerPTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1261 GTCCGGGAGAACCTTACCTCTGTGGGACCTAGAGGTGTGCCAGCCCGGCTGCTCAG 1320
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAenSerMetProArgLeuProThr 460
Db 1321 GGAGAGCCTCAGGCTCTCTGAGAGTCTCTGAGTGGAACTCATGCCCGCTCCCAAG 1380
Qy 461 AspLeuAerValGlyGlyProTrpPheArgHisTyrAerPheArgGlnSerCysTrpVal 480
Db 1381 GACCTGGACGTAGAGGGCCCTTGGTTCGGCATTTGATTTGACAGACAGCTGCTGGGTC 1440
Qy 481 ArgAlaIleSerGlnGlnAerGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 1441 GTGCCCATATCCACGAGAGACCACTGGCCCCCTGTGGCAGGCTGAACACCTGCGAG 1500
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAerSerAerArgGlnGlyThrProPheArg 520
Db 1501 CGGCTGAGATCGCTTTCGCTGCACTGACCACTGATTCGACACAGGACCCCTTCAGA 1560
Qy 521 AlaArgAerGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db 1561 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCCTGCTGCGGCTCCACTTG 1620
Qy 541 GluSerSerGlnPheProGlyPhe 549
Db 1621 GAAGTCTCTCAGTTCCTCCAGGCTTC 1647
```

RESULT 3

AAA15001 standard; cDNA; 1993 BP.

```
XX ID AAA15001 standard; cDNA; 1993 BP.
XX AC AAA15001;
XX DT 21-AUG-2000 (first entry)
XX DE cDNA encoding a human proliferation and apoptosis related protein.
XX KW Human; proliferation and apoptosis related protein; PROAP; psoriasis;
XX KW cell proliferative disorder; immunological disorder; hepatitis;
XX KW reproductive disorder; arteriosclerosis; cirrhosis; melanoma; lymphoma;
XX KW cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia;
XX KW asthma; diabetes mellitus; osteoarthritis; endometriosis;
XX KW uterine fibroid; menstrual cycle; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 42..1691
XX FT /*tag= a
XX FT /product= "proliferation and apoptosis related protein"
XX PN WO200023589-A2.
XX PD 27-APR-2000.
XX PF 19-OCT-1999; 99WO-US024511.
XX PR 20-OCT-1998; 98US-0172216P.
XX PR 04-FEB-1999; 99US-0118559P.
XX PR 11-FEB-1999; 99US-0172229P.
XX PR 22-APR-1999; 99US-0154336P.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lal P;
XX PI Azimzai Y, Baughn MR, Yang J, Shin LL;
XX WPI; 2000-339688/29.
XX P-PSDB; AAY84901.
```

```
XX New human proliferation and apoptosis related protein polypeptides used
PT for diagnosis, treatment and prevention of cell proliferative,
PT immunological and reproductive disorders.
```

Claim 9; Page 117-118; 128pp; English.

The present sequence encodes a human proliferation and apoptosis related protein (PROAP). The polypeptides and polynucleotides can be used for the diagnosis, treatment and prevention of cell proliferative, immunological and reproductive disorders. Disorders associated with decreased expression or activity of include arteriosclerosis, cirrhosis, hepatitis, psoriasis, melanoma, lymphoma and cancers of the breast, brain and prostate, acquired immune deficiency syndrome (AIDS), allergies, anaemia, asthma, diabetes mellitus, osteoarthritis, endometriosis, uterine fibroids and disruptions of the menstrual cycle. Antibodies against PROAP can be used in diagnosis of disorders characterized by PROAP e.g. in ELISA (enzyme linked immunosorbent assays) and the polynucleotides may be used to detect and quantify gene expression in biopsied tissues. These techniques can also be used to monitor regulation of PROAP levels during therapeutic intervention

Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,96e-128	Length:	1993
Score:	2978.00	Matches:	548
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	1
Query Match:	99.8%	Indels:	0
DB:	3	Gaps:	0

US-10-071-838-2 (1-549) x AAA15001 (1-1993)

```
Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAerIleMet 20
Db 42 ATGAGCGTGTAGAGGTGCGGCGAGTTGTGGGCAACAAGCAGAGGACATCATATATG 101
Qy 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGlnLysGlyProLysProPheArg 40
Db 102 AAATACGAAAGAGGACCCAGCTGGGCTCCAGAGACAAAGGGCCTTAAGCTTTTGA 161
Qy 41 SerTyrAsnAsnValAspHisLeuGlyTleValHisGlnThrGlnLeuProLeu 60
Db 162 AGTACAAACAACAACGTCGATCTTGGGATTTGATCATGAGACGGAGCTCCCTCTG 221
Qy 61 ThrAlaArgGluAlaLysGlnIleArgGlnLysSerArgLysSerLysTrpValAsp 80
Db 222 ACTGCCGGGAGCGGAGCAAAATTCGGCGGAGATCAGCCGAAAGACCAAGTGGTGAT 281
Qy 81 MetLeuGlyAspTrpGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 282 ATGCTGGAGACTGGAGAAATACAAAGCAGAAAGCTCATAGATCAGCTTACAG 341
Qy 342 GGAATGCCCATGAACATCCGGGGCCGATGTGTCATCTCTCGAACACTGAGAAATG 401
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
Db 342 GGAATGCCCATGAACATCCGGGGCCGATGTGTCATCTCTCGAACACTGAGAAATG 401
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGln 140
Db 402 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAAGAGAGGCAAGAGCTCTCGAG 461
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 462 CACATCCAGCGCATCGACCGGAGCTAGAGGAGCATTAAGAGACATATTTCTCAGG 521
Qy 161 AspArgTyrGlyThrLysGlnArgGlnLeuHisIleLeuLeuAlaTyrGlnGlnTyr 180
Db 522 GATGATACGGAACCAAGCAGCGGAGACTACTCACAATCTCTCGGATATGAGAGATAT 581
Qy 181 AsnProGlnValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
Db 582 AACCCGAGAGTGCGCTACTGCAAGGAGCTGAGCCATCGCCCGCTTGTCTCTCAT 641
```

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 DB 642 CTTCCTGAGAGATGATCTTGGGCACTGGTGCACCTGCTGGCCATGAGAGCACTCC 701
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 DB 702 CTCGACGAGATTTTCAACAGCCCAATGGCGGAGCGCTCCAGGGGGCTCCAAAGCAACAGAG 761
 QY 241 HisValValAlaThrSerGlnProGlyThrMetGlyHisGlnAspGlyValAspLeuGly 260
 DB 762 CAGTGGTACCGACGTCACCAACCAACATGGGGCATCAGGCAAGAAAGATCTATGT 821
 QY 261 GAG 280
 DB 822 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 881
 QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
 DB 882 CTCACCTGGCGCTGTGGAGCGTGTATCTGGTAGAAGCAAGCGCTTGATCCGATA 941
 QY 301 ThrArgIleAlaPheLeuValGlnGlnLeuArgLeuThrTyrThrSerArgCysGlyPro 320
 DB 942 ACAAGATCCCTTTAAGGTTTCAGACAGACGCTCCAGAAAGATCAGGTGTGCGCCG 1001
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 DB 1002 TGGGACGTTTGTGCAACCGGTTGATGATACCTGGGCGAGAGTAGAGACACTGTGCTC 1061
 QY 341 LysHisLeuArgAlaSerMetClyblyLeuThrArgGlyGlnGlyAspLeuProProPro 360
 DB 1062 AACCATCTTAGGGCTTATGAAAGAACTAACAAGAAAGAGGGGAGCTGCCACCCCA 1121
 QY 361 AlaLeuProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLeu 380
 DB 1122 GCCAAACCCGAGCAAGGAGTGTGCGGATCCAGGCGCTGTGCGGCTTCACTGGCGGGAG 1181
 QY 381 ThrLeuCyblyGlyAspArgGlnAlaProProGluProProAlaArgPheProArgPro 400
 DB 1182 ACCCTCTGCAAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1241
 QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
 DB 1242 ATTTGGTCAAGCTTCCCGGCAAGGGGACCTGTTCTTCCACACCTGCTCTGGTGGGCT 1301
 QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 DB 1302 GTCCGGGAGAGACCTTACCTGTGGGACCTCAGGAGTGTGCGGAGCCGCGCTGGCTCAG 1361
 QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
 DB 1362 GAGGAGCTCAGGGGTTCTGGAGATTCCTGCACTGGAATCCCATGCCCCCTCCCAAG 1421
 QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
 DB 1422 GACCTGAGAGTAAAGGGCCCTTGGTTCCTGCACTTATGATTTCAAGACAGAGCTGCTGCTC 1481
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
 DB 1482 CGGCACTATTCACAGAGGAGACAGAGCTGCGCCCTGCGGAGGCTGAAACCTCGCGAG 1541
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 DB 1542 CGGGTAGATCGGCTTTCCTGCAACCAAGACCTGATTCGACCAAGGAGCAACCCCTTCAG 1601
 QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
 DB 1602 GCTAAGGAGAGAAAGAGCGGTGTGCTCCACCTCAGGGGCTTGCTTGGCGCTCCACTTG 1661
 QY 541 GluSerSerGlnPheProGlyPhe 549
 DB 1662 GAAAGTTCTCAGTTCCTCCAGGCTTC 1688

RESULT 4
 ID ADT88069 standard; cDNA; 1993 BP.
 AC ADT88069;
 DT 30-DEC-2004 (first entry)
 XX
 DE Human proliferation and apoptosis related protein (PROAP)-12 cDNA.
 XX
 KW PROAP; proliferation and apoptosis related protein;
 KW cell proliferative disorder; cancer; atherosclerosis;
 KW immunological disorder; AIDS; acquired immunodeficiency syndrome;
 KW allergy; reproductive disorder; infertility; gene therapy; cytostatic;
 KW antiteriosclerotic; immunosuppressive; anti-HIV; antiallergic;
 KW antifertility; gynaecological; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 42..1691
 FT /tag= a
 FT /product= "Proliferation and apoptosis related protein
 (PROAP)-12"
 XX
 PN US2004203106-A1.
 XX
 PD 14-OCT-2004.
 XX
 PF 05-MAY-2004; 2004US-00839882.
 XX
 PR 19-JAN-1999; 99US-0172216P.
 PR 04-FEB-1999; 99US-0118559P.
 PR 11-FEB-1999; 99US-0172229P.
 PR 22-APR-1999; 99US-0154336P.
 PR 19-OCT-1999; 99WC-0024511.
 PR 11-APR-2001; 2001US-00807452.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Tang YT, Yue H, Hillman JL, Guegler KU, Corley NC, Lai P;
 PI Azimzai Y, Baughn MR, Yang J, Shih LL;
 XX
 DR WPI: 2004-728011/71.
 DR P-PSDB; ADT88050.
 XX
 PT New human proliferation and apoptosis related proteins and
 PT polynucleotides for diagnosing, preventing or treating disorders
 PT associated with aberrant protein expression, e.g. cancer, AIDS,
 PT atherosclerosis or infertility.
 XX
 PS Claim 5; SEQ ID NO 31; 85bp; English.
 XX
 CC The present invention relates to the human proliferation and apoptosis
 CC related protein (PROAP) and its encoding nucleic acid. The invention is
 CC useful for diagnosing, preventing or treating disorders associated with
 CC altered expression or activity of human PROAP, such as cell proliferative
 CC (e.g. cancer or atherosclerosis), immunological (e.g. acquired
 CC immunodeficiency syndrome (AIDS) or allergies) or reproductive (e.g.
 CC infertility) disorders. The invention is also be used in screening for
 CC drugs that may be used for treating or preventing the disorders mentioned
 CC above and in gene therapy. The present sequence is the human PROAP-12
 CC protein encoding cDNA.
 XX
 SQ Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,966-128 Length: 1993
 Score: 2878.00 Matches: 548
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 99.8% Indels: 0
 DB: 13 Gaps: 0

US-10-071-838-2 (1-549) x ADT88069 (1-1993)

Qy 1 MetLeuPValValIGluValAlaGlySerThrTrpAlaGlnGluIleuArgGluAspGlnIleLeuMet 20
 Db 42 ATGAGCGGTGAGAGGTGCGGCGAGTGTGGGCAACAGAGCGAGAGCACTCATTAATG 101
 Qy 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 Db 102 AAATACGAAAGAGGAGCACCGAGCTGGCTGCCAGAGCAAGAGGCGCTTAACCTTTTGA 161
 Qy 41 SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGlyThrGluLeuProProLeu 60
 Db 162 AGCTACAAACAACAGCTGATCTTGGGAGTTGTACATGAGACCGAGCTCCTCTGTG 221
 Qy 61 ThrTrpAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
 Db 222 ACTGCGCGGAGCGCAACCAATTCGCGGAGATCAGCCGAAAGAGCAAGTGGGTGAT 281
 Qy 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
 Db 282 ATGCTGGAGAGCTGGAGAAATACAAAGACAGCAAGAGCTCATAGATCGAGCGTACAG 341
 Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
 Db 342 GGAATGCCCATGAACATCCGGGGCCCGAGTGTGTCAGTCTCTGAAACATGAGGAAATG 401
 Qy 121 LysLeuLysAsnProGlyValArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 Db 402 AAGTTGAAAAACCCCGAAGATACCAATCATGAAGAGAAAGGCAAGAGTCACTGAG 461
 Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 462 CACATCCAGCGGCATCGACCGGAGCGTAAAGCGGACATTAAAGAAACATATTTCTCAG 521
 Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
 Db 522 GATGATACGAGAACCAAGCAGCGGAGACTCTCCCATCTCTGGCATATGAGAGAGAT 581
 Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
 Db 582 AACCGGAGGTGGGTACTGAGGAGCTGAGCCATGCGCGCTTGTCTCTCTCAT 641
 Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGluLeuLeuAlaSerGluIleArgHisSer 220
 Db 642 CTTCTGAGAGAGATGATTCGTGGGCACTGGGCACTGCTGGCAGTGAAGAGCACTCC 701
 Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 Db 702 CTGCAAGGATTTCAACGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGACAGAG 761
 Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysValAspLeuCys 260
 Db 762 CATGTGTAGCCAGCTCAACCCCAAGCAACATGGGCGATCGAGCAACAAAGATCTAAT 821
 Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 822 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGGAATGATTGACGGGATCTCTCTCGG 881
 Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
 Db 882 CTCACCCCTGCGCTGTGGGACGTGTATCTGGTAGAAGCGCAAGGCGTTGATCCGATA 941
 Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 942 ACAAGAAATCGCTTTAAGTTCAAGAGAGCGCTCAGAGAAAGCTCCAGGTGTGGCTG 1001
 Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 1002 TGGGCACTTTTTCGACCGGTTCTTATATCTGGGCGAGGATGAGAACACTGTGCTC 1061
 Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360

Db 1062 AAGCATCTTAGGCGCTTATGAAGAACTAACAAGAAAGAGGGGAGCTTGCCACCCCA 1121
 Qy 361 AlalaProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
 Db 1122 GCCAAACCCAGAGAGGTGTCTGGCATCCAGGCTGTGCGGCTTCACTGGGGGAG 1181
 Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1182 ACCCTTGCAAGGGGAGACAGGACAGGCCCCCTCCAGGCCCAACAGCCCGGTTCCCGGGCCC 1241
 Qy 401 IleTrpSerAlaSerProProAlaArgAlaProArgSerSerThrProCysProGlyValAla 420
 Db 1242 ATTTGTACACTTCCCGGCAAGGAGCTGTTCTTCCACACCTGTCTGTGGGAGCT 1301
 Qy 421 ValArgGluAspTrpThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1302 GTCCGGGAAGACACTTACCTGTGGGACACTCAGGGGTGTGCCAGCCCGGCTTGCTCAG 1361
 Qy 441 GlyLysProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
 Db 1362 GAGAGACCTCAGGGTCTCTGAGATTCCTGCAAGTTCCTGCAAGTTCATGCCCCGCTCCAAAG 1421
 Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
 Db 1422 GACCTGACCTGAGAGGCGCTTGTGTTCCGCAATTATGATTCAGACAGACCTGTGAGTCTC 1481
 Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnIleProAlaGlu 500
 Db 1482 GTGCCATATCCAGAGAGACAGCTGCCCCCTGCTGGAGCTGAACACCTCCGCGAG 1561
 Qy 501 ArgValArgSerAlaPheAlaIleProSerThrAspSerAspGlnGlyThrProPheArg 520
 Db 1542 CGGGTGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGCGACCCCTTCAGA 1601
 Qy 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
 Db 1602 GCTAGGAGCAACAGCGGTGTGCTCCACTCAGAGGCTTGTGCTGTGGCTTCACCTTG 1661
 Qy 541 GluSerSerGlnPheProProGlyPhe 549
 Db 1662 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1688

RESULT 5
 ADC37384
 ID ADC37384 standard; DNA; 1964 BP.
 XX
 AC ADC37384;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 217.
 XX
 KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
 KW immunomodulator; cytotoxic; antimicrobial; Osteopachic; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
 OS Homo sapiens.
 XX
 XX MO2003048202-A2.
 XX
 PD 12-JUN-2003.
 XX
 PP 03-DEC-2002; 2002MO-JP012644.
 XX
 PR 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 PA (ASAH) ASAH KASEI KK.
 XX

PI Matsuda A, Muramatsu S;
XX WPI: 2003-505282/47.
DR P-PSDB; ADC37385.
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
XX useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX ischemic disorders.
PS Claim 4; SEQ ID NO 217; 938bp; English.
XX
XX The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
XX
SQ Sequence 1964 BP; 473 A; 552 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4e-128	Length:	1964
Score:	2975.00	Matches:	548
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	1
Query Match:	99.7%	Indels:	0
DB:	10	Gaps:	0

US-10-071-838-2 (1-549) x ADC37384 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpAlaGlnGluArgLysAspIleMet 20
DB 1 ATGGAAGTGTAAAGTCCGGGCACTTGTGGGCAAGAGCGAGGACATCATTAATG 60
QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 61 AAATACGAAAGGAGACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 120
QY 41 SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
DB 121 AGCTACACAAACAACGTCGATTCATTTGGGATTTGTCATAGACGCGAGCTGCCCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 181 ACTGCGGGAGGCGAAGCAAAATTCGGGGAGATCCAGAAAGCAAGTGGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTrpLys 100
DB 241 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAAAGCTCATAGATCGAGCTACAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnIleMet 120
DB 301 GGAAATCCCATGAACTCCGGGGCCGATGTGTCAGTCTCTCTGAAACATTGAGGAATG 360
QY 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 361 AAGTTAAAAAAGCCCGAAGATACCAAGATCATAGAGGAAGGAGCAAGAGTCAATCGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 421 CACATCAGGCGATCGACCGGAGCGTAAGCGGACATTAAAGAAAGCATTAATCTTCAGG 480
QY 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGlnIleLys 180
DB 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT 540
QY 181 AsnProGluValGlyTrpCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTrp 200
DB 541 AACCCGAGGTGGCTACTGCAAGGAGCTGAGGCACATCCGCTTTGCTCTCTAT 600
QY 201 LeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

DB 601 CTTCCTGAGGAGATGATCTCTGGCACTGTGTCAGCTGCTGGCCAGTAGAGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 661 CTCAGGAGATTTACAGCCCAAAATGGCGGAGCGTCCAGGGGCTCCAAAGCAAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 721 CATGTGTAACCACTGACCAACCAACCAACCATGGGCGCATAGAGCAAGAAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 781 GGGCAGAGTTCCTCCGTTAGGCTGCTCATCCGATATGATTGACGGAGATCTCTCGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValIleTrpLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
DB 841 CTCACCTGGCTGGGACGCTGTGTAATCTGTGTGAAGGCGCAAGCGCTTGATGCCGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 901 ACAGAAATCGCTTTAAGTTTCAGCAAGACCGCTCCAGAAAGCTCCAGGTGTGCCCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 961 TGGGCACTTTTGGCAACCGGTTCTGTGATACCTGGGCGAGGATGAGCACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1021 AAGCATCTTAAGGCTCTTATGAAGAACTAAACAAGAAAGGAGGAGCTCGCAACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgIleGlyLys 380
DB 1081 GCCAAACCGAGAAAGGTCGTGCGATCCAGGCTGTGCGGCTTCACTGCGGGAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1141 ACCCTCTGCAAGGAGGAGACAGGAGCCCTCCAGGCCACAGCCGGTTCGCGGGCC 1200
QY 401 IleTrpSerAlaSerProProAlaGluAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1201 ATTGTGTACGCTTCCCGGACCGGCACTCGTCTTCCACACCTGTCGTGGGGCT 1260
QY 421 ValArgGluAspThrTrpProValGlyTrpGlnGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCCGGAAAGACCTTACCTGTGGCACTCAAGGTGTGCCAGCCCGGCTGGCTCAG 1320
QY 441 GlyLysProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1321 GAGGAGCTCAGGGTTCCTGGAGATTCTCGAGTGAATCCATGCCCCGCTCCAAAG 1380
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTrpAspPheArgGlnSerCysTrpVal 480
DB 1381 GACCTGAGTGAAGGGGCTTGTGTCGCCCATTAATGATTTGAGACAGAGCTGCTGGTTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnAspProAlaGlu 500
DB 1441 CGTCCATATCCAGAGAGACAGCTGCGCCCTGTGGAGGCTAAACACCTTGGGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1501 CGGGTGAGATCGGCTTTCCTGACCCAGACATGATTCGACCAAGGGGACCCCTTCAAG 1560
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1561 GCTAGGAGAGAAACAGCGGTGTCTCCACCTGAGGCGCTTGTCTCTGCGGCTCCACTTG 1620
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1621 GAAAGTTCAGTTCCTCCAGGCTTC 1647

RESULT 6
ABX71403
ID ABX71403 standard; cDNA; 2072 BP.

XX ABX71403;
AC
XX 14-APR-2003 (first entry)
DT
XX Human cell cycle-associated cDNA from clone DKFZp462c3_35p22.
DE
XX Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
KW
XX Homo sapiens.
OS
XX MO200112659-A2.
PN
XX 22-FEB-2001.
PD
XX 18-AUG-2000; 2000MO-IB001496.
PF
XX 18-AUG-1999; 99US-0149499P.
PR
XX 28-SEP-1999; 99US-0156503P.
PR
XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
PA
XX Wiemann S;
PI
XX WPI; 2001-327840/34.
DR
XX P-PSDB; ABUS3234.
DR
XX Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
PT
XX Claim 1; Page 867; 1095pp; English.
PS
XX This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence encodes a polypeptide
CC described in the disclosure of the invention
CC
XX
SQ Sequence 2072 BP; 510 A; 579 C; 586 G; 397 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.5e-127 Length: 2072
Score: 2963.00 Matches: 545
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 5 Indels: 0
DB: Gaps: 0
US-10-071-838-2 (1-549) x ABX71403 (1-2072)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluValArgGluAspIleIleMet 20
DB 99 ATGACGCGTGAAGAGTGGCGGCAAGTGGCGGCAAGAGCCAGAGGACATCATTTATG 158
QY 21 LysTrpGluValGlyValIleArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 159 AAATACGAAAGGAGACACCGAGCTGGCTGCCAGAGGACAAAGGGGCTTAACTTTTGA 218
QY 41 SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGluTrpGluLeuProLeu 60
DB 219 AGCTACCAACAACACGTCGATCATTTGGGATTTGATCACTGAGACGGAGCTCTCTTG 278
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluLysSerArgLysSerLysTrpValAsp 80
DB 279 ACTCCGCGGAGCGGAGCAATTCGCGGAGATCAACCCAAAGAGCAAGTGGGTGAT 338
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerArgLysLeuIleAspArgAlaTyrLys 100

DB 339 ATGCTGGAGACTGGGAAATACAAAGACGAGAAAGCTAGATTCAGCGCTACAG 398
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 399 GGAATGCCATTAACATCCGGGCGCCGATGTGTCATCTCTTCACTGAGGAAATG 458
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 459 AAGTTGAAAAACCCCGAAGATACCATATCATAGAGAAAGGCAAGATCATCTGAG 518
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 519 CACATCCAGCGCATCGACCGGAGCTTAAGCGGACATTAAGAGCACTAATTCCTGAG 578
QY 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 579 GATGATACGAAACCAAGACGCGGAACTACTCATCTCTGTCATATGAGATAC 638
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLysTyr 200
DB 639 AACCCGAGGTGGGCTACTGACAGGACCTGACACATGCGCGCTTCTCTCTAT 698
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluValGHisSer 220
DB 699 CTTCCTGAGGAGGATGCTCTGCGCATCTGCGCATCTGCGCATCTGCGCATCTGCGCATCTC 758
QY 221 LeuGlnGlyPheHisSerProAsnGlyValThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 759 CTGACGGGATTTCCACGCCCAATGGCGGACGCTCCAGGGGCTCCAAAGACCAACAGAG 818
QY 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAspLysLysAspLeuCys 260
DB 819 CATGTGTAGCCAGCTGACAAACCAAGACCATGGGGCATCAGGCAAGAAAGATCTATGT 878
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 879 GGGCAGGTTCCTCCGTAGGCTGCTCATCCGATATGATTGATCGGAGATCTCTCTGCGG 938
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnGlnAlaLeuMetProIle 300
DB 939 CTCACCTTGCGCTGTGGACGTGTATCTGTATGAAAGGCAAGCGGCTGATGCGCATTA 998
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTrpSerArgCysGlyPro 320
DB 999 ACAAGATTCGCTTTAAGTTACGACAGAGCGCTCCAGAAAGCTCAGGTGTGGCCG 1058
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1059 TGGCAGGTTTGGCAACCGGTTCTGTGATACCTGGGCCAGGATGAGACATGTGCTC 1118
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 1119 AAGCATTTAGGGCTTTATGAAAGAACTTAACAAAGAAAGGGGAGACTGCAACCCCA 1178
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
DB 1179 GCCAAACCGAAGAGGAGTGTGCGCATTCAGGCGCTGCGGCTTCACTGGCGGAGAG 1238
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1239 ACCCTCTGCAAGGGGAGCAAGGAGGCTTCCAGGCGCCACACCGCGTTCGCGGCGCC 1298
QY 401 IleTrpSerAlaSerProProAspAlaProArgSerSerThrProCysProGlyValAla 420
DB 1299 ATTGTGACCTTCCCGGACAGGAGCTGTTCTTCCACACCTGTCTCTGTGGGGCT 1358
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1359 GTCCGGGAAACACCTACCTGTGGGACATCAGGGTGTGCCAGCGCGGCGCTGGCTCAG 1418
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1419 GAGAGACCTCAGGGTCTCTGAGATTCCTGAGATTCCTGAGATTCCTGAGATTCCTGAG 1478

QY 461 AspleuapValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
 DB 1479 GACCTGGACGTAGAGGGCCCTGGTTCGCCATTATATTCAGACAGACCTGCTGGCTC 1538
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
 DB 1539 CGTGCCATATCCAGAGAGACAGCTGCCCTCGCTGGCAGGCTGAACCTCTGGAG 1598
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerArgGlnGlyThrProPheArg 520
 DB 1599 CGGGTGGATCGGCTTTCGCTGCACCCAGACATGATCCAGCCAGGCGACCCCTTCAGA 1658
 QY 521 AlaArgAspGlnGluGlyProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
 DB 1659 GCTAGGGACCAACAGACGTGTGCTCCACCTCAGGGCTTGCCTTGGGCTTCACCTTG 1718
 QY 541 GluSerSerGlnPheProGlyPhe 549
 DB 1719 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1745

RESULT 7
 ADQ84596
 ID ADQ84596 standard; cDNA; 2064 BP.
 XX
 AC ADQ84596;
 DT 07-OCT-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #1410.
 XX
 KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KW cancer; cell proliferative disorder; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO2004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003MO-US029126.
 XX
 PR 18-OCT-2002; 2002US-0418988P.
 XX
 PA (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 XX
 PI Wu TD, Zhou Y;
 XX
 DR WPI, 2004-534300/51.
 XX
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 PS Claim 1; SEQ ID NO 1410; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO.1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT)

CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.
 XX
 SQ Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 5,91e-127 Length: 2064
 Score: 2950.00 Matches: 547
 Percent Similarity: 99.3% Conservative: 0
 Best Local Similarity: 99.3% Mismatches: 2
 Query Match: 98.9% Indels: 2
 DB: 13 Gaps: 0

US-10-071-838-2 (1-549) x ADQ84596 (1-2064)

QY 1 MetAspValIaGluValAlaGlySerTyrPAlaGlnGluArgGluAspIleIleMet 20
 DB 99 ATGACGTGTGAGAGGTCCCGGCACTTGTGGGCAACAGAGAGACATCATTTATG 158
 QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 DB 159 AATAAGAAAGGACACCGAGCTGGCTGCAGAGACAAAGGGGCTTAAGCTTTTCA 218
 QY 41 SerTyrAsnAsnAsnValAspHisIleGlyIleValHis--GluThrGluLeuProLys 60
 DB 219 AGCTACAAACAAACATTCATATTTGGATTTGCAAGTAGAGAGAGCTGCTCC 278
 QY 60 eutThraIaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTyrValA 80
 DB 279 TGACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCCGGAAGACCAAGCTGGTGG 338
 QY 80 spMetLeuGluAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrL 100
 DB 339 ATATGCTGGAGACTGGAGAAATACAAAGAGAGCAAGAGCTCATATATTCGATAC 398
 QY 100 ysgIYmetTrometAsnIleArgLysProMetTyrSerValLeuLeuAsnIleGluGluM 120
 DB 399 AGGAGATGCGCATGAAACATCCGGGGCCCATGTGGTCACTCTCCGAAACATGAGAA 458
 QY 120 eTlysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerG 140
 DB 459 TGAAGTTGAAAACCCCGGAAGATACCAATTCAGAGAGAGAGAGAGGCTCATCTG 518
 QY 140 IuhIleIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 160
 DB 519 AGACATCTCAGCGCATCGACCGGAGCTTAAGCGGAGCATTTAAGAAAGCATATATCTTCA 578
 QY 160 rGaAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluT 180
 DB 579 GGAATGATACGGAACCAAGACGGGAACTCTCCACATCTCTCGGATATGAGAGAT 638
 QY 180 yTAsnProGluValGlyTyrCysArgAspLysSerHisIleAlaIleLeuPheLeuLeuT 200

Db 639 ATAAACCGAGTGGGCTACTGACAGGACCTGAGCCACATGCGCGCTTGTTCCTCTCT 698
 Qy YrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnuArgHis 220
 Db 699 ATCTTCTGAGAGAGATGATCTTGGGCACTGGGCACTGGGCACTGGGCACTGGGCACT 758
 Qy 220 erLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnG 240
 Db 759 CCTGACAGGATTTCAACAGCCCAATAGCGGGACCGTCCAGGGGGTCCAAACCAACAGG 818
 Qy 240 LuHLeuValAlaThrSerGlnProGlyThrMetGlyHisGlnAspGlyAspLeuG 260
 Db 819 AGCATGTGTGTGACCACTGACACCAAGACCAAGACCAAGACCAAGACCAAGACCAAG 878
 Qy 260 YSGLYGlnCysSerProLeuGlnCysLeuIleArgIleLeuIleAspGlyIleSerLeuG 280
 Db 879 GTGGGCAAGTGTTCCTGTTAGGCTGCTCATCCGGAATTTAGATGACGGGATCTCTCTCG 938
 Qy 280 YLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnGlnAlaLeuMetProI 300
 Db 939 GGCTCACCTGCTGGCTGTGGGACGTGTATCTGTATGAAAGGGAACAGGCGTTGATGCCGA 998
 Qy 300 IethArgIleAlaPheLeuValGlnGlnIlySerGlyLeuThrLysThrSerArgCysGlyP 320
 Db 999 TAAACAGATGCGCTTTAAGGTTCAGAGAACGCGCTTCACAGAACGCTCCAGGTGTGGCC 1058
 Qy 320 rOTrPALaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValL 340
 Db 1059 CGTGGGCAAGTGTTCCTGTTAGGCTGCTCATCCGGAATTTAGATGACGGGATGAGACCTGTGC 1118
 Qy 340 euLYSHIleuArgAlaSerMetIlyValLeuThrArgIlyGlnGlyAspLeuProProP 360
 Db 1119 TCAGACATCTTAAGGCTCTTATGAAGAACTAACAGAAAGAGGGGACCTGCCACACCC 1178
 Qy 360 roAlaIlyProGluGlnGlySerSerAlaSerArgProValProAlaSerArgIlyGlyL 380
 Db 1179 CAGCCAAACCGAGCAAGGTCGTGCGCATCAGGCTGTCGCGCTTCAAGTGTGGCGGA 1238
 Qy 380 YrThrLeuCysIlyGlnGlyAspArgGlnAlaProProGluIProProAlaArgPheProArgP 400
 Db 1239 AGACCTCTGCAAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1298
 Qy 400 roIleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyA 420
 Db 1299 CCAATTGGTACGTTCCCGCCGACGAGGACCTCGTTCTTCCACACCCGTCTGTGTGGGG 1358
 Qy 420 IaValaIlyGlnAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaG 440
 Db 1359 CTGTGCGGGAAGACACCTACCTGCGGACCTCAGGGGTGCTCCAGCCGCGCTGTGCTC 1418
 Qy 440 IlnGlyIlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProT 460
 Db 1419 AGGAGAGCTCAAGGTTCTGAGATTCCTGCACTGAGAACCTCCATGCCCGCTCCCA 1478
 Qy 460 hrAspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgIlnSerCysTrpV 480
 Db 1479 CGGACCTGGAAGTGAAGGCGCTTGGTTCGCGCATTAATTAATTAATTAATTAATTAATTA 1538
 Qy 480 aLaIArgAlaIleSerGlnGlnuAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaG 500
 Db 1539 TCCCTGTGCATATCCAGAGAGACCAAGCGGCGCTGTGTCAGGCTGACCAACCTGTGGG 1598
 Qy 500 LuAcGValaIlySerSerAlaPheAlaIleProSerThrAspSerAspGlnGlyThrProPheA 520
 Db 1599 AGCGGGTGAATCGGCTTTCCTGACCAACCAAGACCTGATTCGACCAAGGACACCCCTTCA 1658
 Qy 520 rGaIaArgAspGlnGlnProCysAlaIleProThrSerGlyProCysLeuGlyLeuHisI 540
 Db 1659 GAGCTAGGAGACCAAGCGGTGTGTCTCCACTCAGGCGCTTGTGCTGTGGGCGCTCCACT 1718
 Qy 540 euGlnSerSerGlnPheProProGlyPhe 549
 Db 1719 TGGAAAGTTCAGATTCCTCCAGGCTTC 1747

RESULT 8
 ADQ6595
 ID ADQ6595 standard; cDNA; 2064 BP.
 AC ADQ6595;
 DT 07-OCT-2004 (first entry)
 XX Human tumour-associated antigenic target (TAT) cDNA sequence #3468.
 XX human; tumour-associated antigenic target; TAT; cytotoxic; gene therapy;
 KM cancer; cell proliferative disorder; gene; ss.
 XX Homo sapiens.
 OS
 PN W02004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003MO-US029126.
 XX
 FR 18-OCT-2002; 2002US-0418988P.
 XX
 PA (GERTH) GENENTECH INC.
 PA (WU) D. WU T D.
 PA (ZHOU) ZHOU Y.
 XX
 PI Wu TD, Zhou Y;
 XX
 DR WPI; 2004-534300/51.
 XX
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 preventing or treating cell proliferative disorders such as cancer.
 XX
 XX Claim 1; SEQ ID NO 3468; 5504bp; English.
 XX
 CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO.1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potential effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a

CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,91e-127	Length:	2064
Score:	2950.00	Matches:	547
Percent Similarity:	99.3%	Conservative:	0
Best Local Similarity:	99.3%	Mismatches:	2
Query Match:	98.9%	Indels:	2
DB:	13	Gaps:	0

US-10-071-838-2 (1-549) x ADQ86595 (1-2064)

```

QY      1 MetAapValValGluValAlaGlySerTTPAlaGlnGluArgGluAapPleIleMet 20
DB      99 ATGACCTGGTAGAGTCTCGCGGCGCATGTGGGCAAGAGCGAGCATCATATATG 158
QY      21 LysETGluLysGlyHisArgAlaGlyLeuProGluAapLysGlyProLysProPhaArg 40
DB      159 AAATACAAAGGAGCACCGAGCTGGGCTCCAGAGGACAAAGGGGCTTAAGCTTTTCGA 218
QY      41 SerTyrAapAapAapValAapHisLeuGlyIleValHis--GluThrGluLeuProProl 60
DB      219 AGCTACACAAACAACGTCGATCATTGGGATTTGTACAAGTGAACGAGCTGCTCTC 278
QY      60 eutThAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTTPValA 80
DB      279 TGACTGGCGGAGGCGGACCAATTCGCGGAGATCACCGCAAGAGGCAAGTGGGTGG 338
QY      80 sptMetLeuGlyAapTTPGluLysTyrLysSerSerArgLysLeuIleAapArgAlaTyrL 100
DB      339 ATATGCTGGGAGACTGGGAAATATCAAAAGCAGACAAAGCTCATATGATCGAGCTGACA 338
QY      100 YagLysMetProMetAapIleArgGlyProMetTTPSerValLeuLeuAapIleGluGlu 120
DB      399 AGGGAATGCCCATGAAACATCCGGGGCCCGATGTGTGATGCTCTCTGAAACCTGAGAAA 458
QY      120 eLysLeuLysAapProGlyValArgTyrGlnIleMetLysGluLysGlyLysArgSerSerg 140
DB      459 TGAAGTTGAAAACCCCGGAAAGATACCAAGATCATGAAGAGGAGGCAAGGCTCATCTG 518
QY      140 IuhIstIleGlnArgGlyLeaAapArgAapValSerGlyThrLeuArgLysHisIlePhePheA 160
DB      519 AGACACATCCAGCGCATCGACCGGAGCTGAAGCGGACATTAGGAAACATATATTTCTCA 578
QY      160 rgaAapArgTyrGlyThrLysGlnArgGlyLeuLeuHisIleLeuLeuAlaTyrGluGlu 180
DB      579 GGGATCGATCGGAACCAAGACGCGGAACTTATCCCATCTCTCTGCAATGAGGAGT 638
QY      180 YrAapProGluValGlyTyrCysArgAapLysSerHisIleAlaLeuPheLeuLeuT 200
DB      639 ATAAACCGAGAGTGGGCTACTGCAAGGAGCTGAGCCCATCTGCGCTTCTCTCTCTCT 698
QY      200 YrLeuProGluGluAapAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHis 220
DB      699 ATCTCTCTGAGAGATGATCTTCTGGGCACTGTGCGATCTGCGGCAATGAGAGGACT 758
QY      220 eLysGlnGlyPheHisSerProAapGlyGlyTyrThrValGlnGlyLysGlnAapGlnGln 240
DB      759 CCTCGAGGAGTTTCACAGCCCAATGGCGGAGCTTCAGGGGCTTCACAGACCAACAGG 818
QY      240 IuhIstValValAalThrSerGlnProLysThrMetGlyHisGlnAapLysLysAapLeu 260
DB      819 AGCATGGTGAACACGTCACAAACCAAGCATGGGCGATCGAGCAAGAAAGATCTAT 878
QY      260 YagLysGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAapGlyIleSerLeu 280
DB      879 GTGGGAGTGTCCCGTATAGCTGCTCATCCGGAATATGACGGGATCTCTCTCG 938
QY      280 YLeuThrLeuArgLeuTTPAapValTyrLeuValGluGlyGlnAlaLeuMetProI 300

```

```

DB      939 GGCTACCCCTGGCGCTGTGGAGCTGTATCTGGTAAGAGGCAACAGCGCTTATGCCGA 998
QY      300 IeThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyP 320
DB      999 TAAACAAGATCGCTTTTAAGGTTTACAGAGAGAGCCCTCAAGAGACCTCCAGGCTGTGGCC 1058
QY      320 rOTTPAlaArgPheCysAapAapArgPheValAapThrTTPAlaArgAapGluAapThrVal 340
DB      1059 CGTGGGACATTTTTCACACCGGTTCTTTATACCTGGGCGAGGAGATGAGACATGTGC 1118
QY      340 eulYrHisLeuAArgAlaSerMetLysLysLeuThrArgLysGlnGlyAapLysProProl 360
DB      1119 TCAAGATCTTAAGGCTCTTATGAAGAACTTAAGAAAGAGAGGGGAGCTTGCACATCCC 1178
QY      360 roAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyL 380
DB      1179 CAGCCAAACCCGAGCAAGGCTCTCGCATCCAGGCTGTGGCGGCTTCACGTGGCGGGA 1238
QY      380 YrThrLeuCysLysGlyAapArgGlnAlaProProGlyProProAlaArgPheProAap 400
DB      1239 AGACCTCTGCAAGGGGAGAGGAGGAGCCCTCCAGGACCAAGAGCCGAGTCCCGGCGC 1298
QY      400 rOileTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyA 420
DB      1299 CATTTGTGATGAGCTTCCCGGCAAGGAGCTGTTCTTCCACACCTGTCTGTGGGG 1358
QY      420 IeValArgGluAapThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaG 440
DB      1359 CTGTCCGGGAGACACCTACCTGAGGAGCTCAGAGGTGTGCCAGGCCCGCTGGCTC 1418
QY      440 IngLysLysProGlnGlySerTTPArgPheLeuGlnThrAapSerMetProAapLeuPro 460
DB      1419 AGGAGAGACTCAGGGTCTGTGAGATCTCTGAGTGAATCTCATAGCCCGGCTCCAA 1478
QY      460 hrAapLeuAapValGluGlyProTTPPheArgHisTyrAapPheAArgGlnSerCysTTP 480
DB      1479 CGGAGCTGAGCGTGAAGGGGCTTGTGTCGCAATATGATTTCAAGAGAGCTGTGGG 1538
QY      480 aIArgAlaIleSerGlnGluAapGlnLeuAlaProCysTTPGlnAlaGlnHisProAlaG 500
DB      1539 TCCGTCCATATCCAGAGAGACCACTGGCCCCCTGTGGCAGGCTGAACACCTGGCGG 1598
QY      500 IuhArgValArgSerAlaPheAlaAlaProSerThrAapSerAapGlnGlyThrProPhe 520
DB      1599 AGCGGATGAGATCGGCTTTCGTGCAACCAAGCATATTCGACAGGGGACCCCTTCA 1658
QY      520 rgaIaArgAapGluGlnProCysAlaProThrSerGlyProCysLysGlyLeuHisIle 540
DB      1659 GAGCTAGGAGCAAGAGCGGTGTGCTCCACCTCAAGGCTTGTCTGTGGCTTCACT 1718
QY      540 eugLysSerGlnPheProProGlyPhe 549
DB      1719 TGGAAAGTTCTCAGTTCCTCCAGGCTTC 1747

```

RESULT 9

ABSS4707 standard; cDNA; 2146 BP.

XX ABSS4707;

XX 28-NOV-2002 (first entry)

XX cDNA encoding human PRC17 protein splice variant 1.

XX Human; PRC17; prostate cancer; ovarian cancer; lung cancer;

XX breast cancer; cytostatic; chromosome 17q11-12; gene; ss; splice variant.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS

FT /*tag= a /product= "Human PRC17 protein splice variant 1"

/transl except= (pos:421..422, aa:Asp)
/note= "This codon has an apparent 1 nucleotide deletion
which alters the reading frame"

XX WO200262958-A2.

XX 15-AUG-2002.

XX 08-FEB-2002; 2002WO-US003457.

XX 08-FEB-2001; 2001US-0267615P.

XX (TULIA-) TULIARIK INC.

XX LI J, Powers S, Xiang P, Peng Y;

XX WPI; 2002-706902/76.

XX P-PSDB; ABG70737.

XX Novel isolated PRC17 polypeptide useful diagnostically or prognostically

XX PT or detect diseases or conditions associated with altered PRC17 activity

XX PS Claim 24; Page 63-64; 78pp; English.

XX The present invention relates to a new PRC17 polypeptide. The invention

XX is useful for detecting cancer cells (such as prostate tissue, breast

XX tissue, lung tissue, ovarian tissue) in a biological sample. The

XX invention is further useful for monitoring the efficacy of a therapeutic

XX treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,

XX breast cancer). The molecules of the invention are useful diagnostically

XX or prognostically to detect diseases or conditions associated with

XX altered PRC17 activity or expression relative to normal, for example

XX cancer. The present nucleic acid sequence encodes the human PRC17 protein

XX splice variant 1 of the invention. The human PRC17 gene is located on

XX chromosome 17q11-12

XX SQ Sequence 2146 BP; 494 A; 619 C; 616 G; 417 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

5,49e-125
2907.50
89.7%
97.5%
6Length:
Matches:
Conservative:
Mismatches:
Indels:2146
547
0
2
62

Gaps: 1

US-10-071-838-2 (1-549) x ABSS4707 (1-2146)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleuMet 20

Db 1 ATGACGCTGGTGAAGGTGCGGCGACGTTGGTGGCAAGACGAGACATCATTTATG 60

QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40

Db 61 AAATACGAAAGGAGCACCGAGCTGGCTGCCAGAGGACAAAGGGGCTTAACCTTTTGA 120

QY 41 SerTyrAsnAsnValAspHisIleuGlyIleValHis----- 53

Db 121 AGCTAACACAAACACGTCGATCTTGGGGAATTGAAGTCTCTGCGCTCTGGAGTCA 180

QY 53 ----- 53

Db 181 GCCCACAGAAAGCCCTTGTCTCTCCCTGTGCTTCTCTCTGGGCTGAGCCCTGAG 240

QY 53 ----- 53

Db 241 CTGGAAGAGGACAGAGCAATCTTTCTGGGGGTGCGACCAAGGCTGGGCGGCTCCAG 300

QY 54 -----GluThrGluLeuProPro 59

Db 301 GCCCGTGCAGTCTCTCAGCTCTGCTGGCTGGTTCCTTAAGTGAAGACGAGCTGCTCT 360

QY 60 LeuThrAlaArgLysAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpVal 79

Db 361 CTGACCTCGGGGAGGAGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGACAGTGGGT- 419

QY 80 AspMetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyr 99

Db 420 GATATGCTGGGAGACTGGGAGAAATACAAAAGCAGGAAAGCTCATAGATCGGCGTAC 479

QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlu 119

Db 480 AAGGAAATGCCATGAACATCCGGGCGCGATGTGTCACTCTCTGAACTTGAAGAA 539

QY 120 MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSer 139

Db 540 ATGAAGTTGAATAAAACCCCGGAAGATACAGATCATGAAGGAAAGGCAAGAGTCACT 599

QY 140 GluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 159

Db 600 GAGCAGATCCAGCCCATCGACCCGGAGCTTAAGCGGACATTAAGGAAGCATATATCTTC 659

QY 160 ArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGlu 179

Db 660 AGGATGATACGGAACCAAGCAGCGGAACTACTCAATCTCTCTGCAATAGAGAG 719

QY 180 TyrAsnProGluValGlyTyrCysArgAspLysSerHisIleAlaLeuPheLeuLeu 199

Db 720 TATTAACCGGAGGTGGGCTACTGACGGACCTGACCAATCCGCGCTTGTCTCTCTC 779

QY 200 TyrLeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 219

Db 780 TATCTTCTGAGGGAGGTGATCTGTGGCACTGTGTGACGTGCGGCAAGTGAAGGACAC 839

QY 220 SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 239

Db 840 TCCCTGAGGAGATTTCAAGCCCAATAGCGGAGCCCTCCAGGGGCTTCAAGACCAACAG 899

QY 240 GluHisValAlaAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 259

Db 900 GAGCATGTGGTAGCAGTCATCAACCAAGCATTGGGGCTACGACAAAGAAATCTA 959

QY 260 CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu 279

Db 960 TGTGGGAGAGTTCCTCCGTTAGGCTGCTCATCCGGAATATTAATGACGGGATCTCTC 1019

QY 280 GlyLeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnGlnAlaLeuMetPro 299

Db 1020 GGGCTCACCTGCGCCTGTGGAGCTTATCTGTGAAGACCAAGCGGCTTGATCCG 1079

QY 300 IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTrpSerArgCysGly 319

Db 1080 ATTAACAAGAAATCGCTTTAAGTTACAGCAGAAGGCTTCAAGAAAGCTCCAGGTGGC 1139

QY 320 ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 339

Db 1140 CCGTGGGACAGTTTTCCAACCGGTTGTGATACCTCGGGCCAGGGATGAGACACTGTG 1199

QY 340 LeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnLysAspLeuProPro 359

Db 1200 CTCAAGCATCTTAAGGCTCTTATGAAGAACTTAACAAGAAAGGAGGAGACTCCGACCC 1259

QY 360 ProAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGly 379

Db 1260 CAGCCAAACCCAGAGCAAGGGGTGTGGCATTCAGGCTGTGGCGGCTTACGCGGGGG 1319

QY 380 LysThrLeuCysLysGlyAspArgGlnAlaProProGlyLysProAlaArgPheProArg 399

Db 1320 AAGACCTCTGCAAGGGGAGCAGGCAAGCCCTTCAAGGCCACACAGCCCGGCTCCCGG 1379

QY 400 ProIleThrSerLysSerProProArgAlaProArgSerSerThrProCysProGlyGly 419

Db 1380 CCCATTGTGTGAGTTCCTCCGCGCACCGGCACTCTGTCTTCCACACCTTGTCCGTGG 1439

QY 420 AlaValArgLysAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAla 439

```

Db      1440 GCTGTCGGGAGAGACCTACCTGTGGGACATCAGGGTGTGCCAGCCCGGCTGCT 1499
Qy      440 GINGLYGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuPro 459
Db      1500 CAGGAGGAGCCTCAGGGTTCCTGGAGATTCTGCACTGCAACCTCCATGCCCGCCCTCCCA 1559
Qy      460 ThrAspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrp 479
Db      1560 ACGGACCTGACGTAGAGGGCCCTTGTGCTCCGCAATTATGATTTTCAGACAGAGCTGTGG 1619
Qy      480 ValArgAlaIleSerGlnGlnLeuAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAla 499
Db      1620 GTCCGTCGATATCCAGAGAGGACGAGCTGGCCCTGTGGAGAGGCTAACACCTGTGG 1679
Qy      500 GlnArgValArgSerAlaPheAlaAlaProSerThrArgSerAspGlnGlyThrProPhe 519
Db      1680 GAGCGGTGAGATCGGCTTTCCTGCTGCACCCACACTGATTCGACACGAGGCAACCCCTTC 1739
Qy      520 ArgAlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHis 539
Db      1740 AGAGCTAGGAGACGACAGCCGTGTGCTCCACCTCAGGAGCTTGCTGTGCGGCTCCAC 1799
Qy      540 LeuGlnSerSerGlnPheProProGlyPhe 549
Db      1800 TTGAAAAGTTCTCAGTTCCCTCCAGGCTTC 1829

```

RESULT 10

```

ADCC37380
ID      ADC37380 standard; DNA; 2304 BP.
XX
AC      ADC37380;

```

```

XX      18-DEC-2003 (first entry)
DE      Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 213.
KW      Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW      cancer; infectious disease; bone disease; AIDS;
KW      neurodegenerative disease; ischemic disorder; Antiinflammatory;
KW      immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW      Neuroprotective; Noctropic; Cardiant; Gene therapy; human; gene; ds.
OS      Homo sapiens.
XX
FN      WO2003048202-A2.
XX      12-JUN-2003.
XX      03-DEC-2002; 2002WO-JP012644.
XX      03-DEC-2001; 2001JP-00368692.
XX      05-DEC-2001; 2001US-0335829P.
XX      03-OCT-2002; 2002JP-00291302.
XX      04-OCT-2002; 2002US-0415763P.
XX      (ASAH ) ASAH KASEI KK.
XX      Matsuda A, Muramatsu S;
XX      MPI; 2003-505282/47.
XX      P-PSDB; ADC37381.
XX      New purified protein that activates nuclear factor kappa B (NF-kappaB),
XX      useful for treating inflammation, autoimmune diseases, cancers,
XX      infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX      ischemic disorders.
XX      Claim 4; SEQ ID NO 213; 938pp; English.
XX
CC      The present invention relates to novel proteins and their coding
CC      sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC      kappaB). The proteins and their coding sequences are useful for treating

```

CC a disease associated with NF-kappaB activation, such as inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases, or ischemic disorders.

XX Sequence 2304 BP; 569 A; 610 C; 636 G; 489 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.:	Length:	Matches:	Conservative:
Score:	6,526-125	2304	540
Percent Similarity:	2906.50	98.4%	0
Best Local Similarity:	98.4%	Mismatches:	2
Query Match:	97.4%	Indels:	8
DB:	10	Gaps:	1

US-10-071-838-2 (1-549) x ADC37380 (1-2304)

```

Qy      1 MetAspValValGlnValAlaGlySerTrpAlaGlnGlnArgGlnAspIleIleMet 20
Db      362 ATGACGTGTGAGAGTGGCGCGGAGTTGTGGCACAAGACGAGAGCATATTATG 421
Qy      21 LysTyrGlnGlyGlyHisArgAlaGlyLeuProGlnAspGlyGlyProLysProPheArg 40
Db      422 AAATACGAAAAGGAGCACCGAGCTGGCTGCCAGAGGACAGAGGGGCTTAAGCTTTTCCA 481
Qy      41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGlnThrGlnLeuProPheLeu 60
Db      482 AGCTACAAACAAACGTCGATATTGGGGATTTGATACATGACAGAGAGCTGCTCTCTCG 541
Qy      61 ThrAlaArgGlnAlaGlyGlnIleArgArgGlnIleSerArgGlySerTrpValAsp 80
Db      542 ACTGCGCGGAGGCGCAAGCAAAATTCGCGGAGATCACCGAAAGCAAGAGTGGTGAATG 601
Qy      81 MetLeuGlyAspTrpGlnGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db      602 ATGCTGGAGACTGGGAAATAATACAAAGCAGCAAAAGCTATGATGAGCTGACAAAG 661
Qy      101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGlnMet 120
Db      662 GGAATGCCCATGAAATATCCGGGGCCGAGTGTGTAGTCTCTTAAACATGAGGAATATG 721
Qy      121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnGlyLysArgSerSerGln 140
Db      722 AAGTTGAAAACCCCGGAAAGATACCAATCATGAAAGAAAGGCAAGATCATCTGAG 781
Qy      141 HisIleGlnArgIleAspArgAspValSerGlyThrIleAspGlyHisIlePhePheArg 160
Db      782 CACATCCAGCCCATGACCGGGAGCGTAAGCGGACATTAAGAGCATATATTTCTTCAGG 841
Qy      161 AspArgTyrGlyThrLysGlnArgGlnLeuLeuHisIleLeuLeuAlaTyrGlnGlnLys 180
Db      842 GATCGATACGGAACCAACAGCGGAGAACTACATCCATCTCTGGCATATGAGAGATAT 901
Qy      181 AsnProGlnValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db      902 AACCCGAGGTGGGCTATCGAGGAGCTGAGCCACATCGCGGCTTGTCTCTCTATAT 961
Qy      201 LeuProGlnGlnAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnArgHisSer 220
Db      962 CTTCCTGAGAGAGATGATCTTGAGGCACTGGTGAGCTGTGGCAGAGAGAGGACATCC 1021
Qy      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlnGlnAspGlnGlnGln 240
Db      1022 CTGCAGGGATTTTCAAGCCAAATGGCGGAGCCGTCCGAGGGGCTTCCAAGACCAAGAG 1081
Qy      241 HisValValAlaTrnSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db      1082 CATGTGTAGCCACGTCAACACCAAGACATGGGGCATCAGAGCAAAAGATCTATGT 1141
Qy      261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db      1142 GGGCAGTGTCTCCGTGTAGGCTGCTCATCCGGAATATGACGGGATCTCTCGGG 1201
Qy      281 LeuThrLeuAspLeuTrpAspValTyrLeuValGlnGlnGlnAlaLeuMetProIle 300

```

```
Db 1202 CTCACCTGGCGCTGTGGAGCGTGTATCTGTAGAAAGGCAAGGGCTGATGCCGATA 1261
QY 301 ThTArG11eAlaPhelysValGInGlnIuYArGleuThrYsThrsrArGySgIyPro 320
Db 1262 ACAAGATCCGCTTTAAGTTCAAGCAGAACCGCTCCAGAAAGCTCAGAGTGTGGCCG 1321
QY 321 TrpAlaArpPheCyAenaArpPheValAspThrTrpAlaArpAspGluAspThrValLeu 340
Db 1322 TGGGCAAGTTTTCGAAACCGGTTCTGTATACCTGGGCCAGGATGAGACACTGTCTC 1381
QY 341 LysHisIleuArGAlaSerMetLysLysLeuThrArgIleGlnIuYAspLeuProPro 360
Db 1382 AAGCATCTTAGGGCGCTTATAGAAACTAACAAGAAAGAGGGGAGCCTGCCACCCCA 1441
QY 361 AlaIysProGluGlnIuYSerSerAlaSerArpProValProAlaSerArGyIyLys 380
Db 1442 GCCAAACCCAGAGAGGGTGTGTGGCATCCAGGCTGTGTGCCGCTTCACTGGCGGGAG 1501
QY 381 ThrLeuCyAlysGlyAspArpGlnAlaProProGlyProProAlaArpPheProArGPro 400
Db 1502 ACCCTCTGCAAGGGGAGCAGGAGCCCTCCAGGC-----CCA 1540
QY 401 IleTrpSerAlaSerProProArGAlaProArGSerSerThrProCySProGlyAla 420
Db 1541 TT-TGGTCACTTCCCGCCAGCGGCACTCGTTCTTCCACACCTGTCTCGTGGGGGCT 1599
QY 421 ValArGluAspThrYrProValGlyThGlnGlyValProSerProAlaIleuAlaGln 440
Db 1600 GTCCGGGAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGGCTCAG 1659
QY 441 G1yG1yProGlnIuYSerTrpArpPheLeuGlnTrpAsnSerMetProArGLeuProThr 460
Db 1660 GAGAGACCTCAGGGTCTCTGAGATTCTCTGACATGGAACTCCATGCCCGCTCCCAAG 1719
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTrpAspPheArgIuSerCySTrpVal 480
Db 1720 GACCTGAGCTAGAGGGGCCCTGTGGTTCGGCCATTATGATTTCAAGACAGACTGTGGGTC 1779
QY 481 ArgAlaIleSerGlnIuYAspGlnLeuAlaProCySTrpGlnAlaGluIleProAlaGlu 500
Db 1780 CGTGCCTATCTCCAGGAGGACAGCTGCCCTCGCTGGAGGCTGAACACCTGTCCGAG 1839
QY 501 ArgValArpSerAlaPheAlaAlaProSerThrAspSerAspGlnIuYTrpProPheArg 520
Db 1840 CGGGTGAATCGGCTTTCGCTGCACCAAGCACTATTCCAGCCAGGGCACCCTTACAG 1899
QY 521 AlaArpAspGluGlnProCySAlaProThrSerGlyProCySLeuCySgIyLeuHisLeu 540
Db 1900 GCTAGGAGCAAGACGCGTGTGCTCCACCTCAGAGGCTTGCTGTGGGCTTCACTTG 1959
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1960 GAAAGTTCTCAGTTCTCTCCAGGCTTC 1986

RESULT 11
ADc37378 standard; DNA; 2304 BP.
ID ADc37378
AC ADc37378;
XX
DT 18-DEC-2003 (first entry)
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 211.
XX
XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
XX cancer; infectious disease; bone disease; AIDS;
XX neurodegenerative disease; ischaemic disorder; Anti-inflammatory;
XX immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
XX Neuroprotective; Neotropic; Cardiac; Gene therapy; human; gene; ds.
XX
XX Homo sapiens.
XX
```

```
PN WO2003048202-A2.
XX
XX 12-JUN-2003.
PD
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
XX 03-DEC-2001; 2001JP-0036692.
PR 05-DEC-2001; 2001US-035829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
XX (ASAH ) ASAH KASEI KK.
XX
XX Matsuda A, Muramatsu S;
XX
XX WPI: 2003-505282/47.
XX
XX P-PSDB; ADc37379.
XX
XX DR
XX
XX DR
XX
XX PT New purified protein that activates nuclear factor kappa B (NF-kappaB) ,
XX PT useful for treating inflammation, autoimmune diseases, cancers,
XX PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX PT ischemic disorders.
XX
XX Claim 4; SEQ ID NO 211; 938pp; English.
XX
XX The present invention relates to novel proteins and their coding
XX sequences (ADc37168-ADc37455), which activate nuclear factor kappa B (NF-
XX kappaB). The proteins and their coding sequences are useful for treating
XX a disease associated with NF-kappaB activation, such as inflammation,
XX autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
XX neurodegenerative diseases, or ischemic disorders.
XX
XX SQ Sequence 2304 BP; 570 A; 609 C; 636 G; 489 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,52e-124 Length: 2304
Score: 2898.50 Matches: 539
Percent Similarity: 98.2% Conservative: 0
Best Local Similarity: 98.2% Mismatches: 3
Query Match: 97.2% Indels: 8
DB: 10 Gaps: 1

US-10-071-838-2 (1-549) x ADc37378 (1-2304)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnIuYAspIleIleMet 20
Db 362 ATGAGCTGTGTAGAGGTCCGGCACTTGTGGGCAAGAGGAGACATCATTTATG 421
QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspIySgIyProLysProPheArg 40
Db 422 AAATACGAAAGGAGCACCGAGCTGTGGCTGCAGAGGACAGGGGCTTAAGCCTTTTCCG 481
QY 41 SerTrpAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 482 AGCTACAAACAACATCGATCTTGTGGGATTGTACATGACGAGAGCTGCTCTCTG 541
QY 61 ThrAlaArgGluAlaLysGlnIleArGArGluIleSerArGySeriYsTrpValAsp 80
Db 542 ACTGCCGGGAGGCGAAGCAATTCGCGCGGAGATCAGCCGAAAGAGCAAGTGGTGGAT 601
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArGyLysLeuIleAspArgAlaTrpLys 100
Db 602 ATGCTGGAGACTGGGAGAAATCAAAAGCAGACGAAAGCTCTTAATCGAGCTGCAAG 661
QY 101 GlyMetProMetLeuIleArgIyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 662 GGAATGCCCATGAAACATCCGGGGCCGATGTGTGCTCTCTTGAACACTGAGAAATG 721
QY 121 LysLeuLysAspProGlyArgTrpGlnIleMetLysGluLysGlyLysArGSerSerGlu 140
Db 722 AAGTTAAAAACCCCGAAGATTCAGATCATTAAGAGAGGAGGCAAGGTATCTGAG 781
QY 141 HisIleGlnArgIleAspArgValSerGlyThrLeuArgLysHisIlePhePheArg 160
```

Db 782 CACATCCAGCGCATCCAGCGGACGTAAAGGGGACATTAAAGAGCATATATCTTCAGG 841
Qy 161 AsparGlyTgLYThrLySGlnArgGlnLeuLeuHisLeuLeuAlaTyrGlyGluTyr 180
Db 842 GATCGATACGGAACCAAGACGCGGAACTACTCCACATCTCTCCGGCATATGAGGAGTAT 901
Qy 181 AspProGluValGlyTyrCybArgAspLeuSerHisLeuAlaLeuPheLeuLeuTyr 200
Db 902 AACCCGAGGTGGGCTTCTGAGGGACCTGAGCCACATCGCGCTTGTTCTCTCTAT 961
Qy 201 LeuProGluGlnAspAlaPheTyrPalaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 962 CTTCTGAGGAGGATGATCTGCGCATCTGGGCACTGGTGACGTGCTGGCATGAGAGCACTCC 1021
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlu 240
Db 1022 CTCGAGGATTTCCACGCCAAATGCGGGACCGTCCAGGGGCTCCAAAGACCAACAGAG 1081
Qy 241 HisValValAlaThrSerGlnProLyThrMerGlyHisGlnAspLybAspLeuCyb 260
Db 1082 CATGTGTAGCCACGTACAAACCAAGACATGGGGCATGACGAAAGATCTATGT 1141
Qy 261 GlyGlnCySerProLeuGlyCybLeuLeuArgLeuLeuAspGlyLeuSerLeuGly 280
Db 1142 GGGCAGGTTCCTCCGTTAGGCTGCTCATCCGATATTGATGACGGGATCTCTCGGG 1201
Qy 281 LeuThrLeuArgLeuTyrAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 1202 CTCACCTCGCGCTGTGGAGCGTGTATCTGTAGAGCGAACGCGGTGATGCCGATA 1261
Qy 301 ThrArgGlyLeuAlaPheLybValGlnGlnLybArgLeuThrLybThrSerArgCybGlyPro 320
Db 1262 ACAAGATTCCTCTTAAAGTTCAACAGAACGCGCTCAAGAACGTCACAGTGTGGCCG 1321
Qy 321 TrpAlaArgPheCybAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1322 TGGGCACTTTTGGCAACCGGTTCTGTGATACCTGGGCCAGGATGAGACATCTGCTC 1381
Qy 341 LybHisLeuArgAlaSerMetLybLybLeuThrArgLybGlnGlyAspLeuProProPro 360
Db 1382 AAGCATCTTAAAGGCTCTATGAAGAACTAAACAAAGAGGAGGAGCTGCAACCCCA 1441
Qy 361 AlaLybProGlnGlnGlySerSerAlaSerArgProValProAlaSerAlaGlyLyb 380
Db 1442 GCCAAACCCGAGAGGGGTGTGGCATTCAGGCTGTGCGGCTTCACTGCGGGAG 1501
Qy 381 ThrLeuCybLybGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1502 ACCCTCTGCAAGGGGAGAGGAGGCGCCCTCCAGGC-----CCA 1540
Qy 401 IleTrpSerAlaSerProArgAlaProArgSerSerThrProCybProGlyGlyAla 420
Db 1541 TT-TGGTCACTTCCCGGCAAGGAGCACTGTTCTTCCACACCTGTCTCTGGGGGCT 1599
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1600 GTCCGGGAAGACCTAACCTGTGGGACCTCAGAGGTGTGCCAGCCGCGCTGCGTCAG 1659
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1660 GAGAGACCTCAGGCTCTGAGATTCCTGAGAACTCCATGCCCGCCCTCCAAAG 1719
Qy 461 AspLeuAspValGlnGlyProTyrPheArgHisTyrAspPheArgGlnSerCybTrpVal 480
Db 1720 GACCTGAGCGTAAAGGCGCTTGGTTCCGCCATTATGATTTCAGACAGAGCTCGGCTC 1779
Qy 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCybTrpGlnAlaGluHisProAlaGlu 500
Db 1780 CGGCAATATCCAGAGAGCAAGCTGGGCGCCCTGCTGGAGGCTGAAACCTCGGGAG 1839
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520

Db 1840 CGGGTAGATCGGCTTTCGTGACACCGACACTGATTCGAGCAACGAGGACCCCTTCAGA 1899
Qy 521 AlaArgAspGlnGlnProCybAlaProThrSerGlyProCybLeuCybGlyLeuHisLeu 540
Db 1900 GCTAGGAGCAACAGCGTGTGCTCCCACTCAGAGGCGCTTGGCTGCGCTCCACTTG 1959
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 1960 GAAAGTTCTGATTCCTCCAGGCTTC 1986
RESULT 12
ID ABZ11810
ABZ11810 standard; cDNA; 2084 BP.
XX
AC ABZ11810;
XX
XX 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 692.
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX cell-proliferative disorder; neurodegenerative disease; bacterial;
XX Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX arthritis; cytostatic; immunomodulator; noctropic; neuroprotective;
XX antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
XX haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
XX antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
XX
XX Homo sapiens.
XX
XX NO200270539-A2.
XX
PD 12-SEP-2002.
XX
XX
XX 05-MAR-2002; 2002MO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Zhou P, Goodrich RW, Agungi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX Wehrman T, Wang J, Wang D, Drmanac RT;
XX
XX MPI; 2002-759812/82.
XX P-PSDB; ABP65933.
XX
XX New polynucleotides comprising sequences assembled from expressed
XX PT sequence tags (ESTs), useful for treating cell-proliferative,
XX PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
XX or coagulation disorders.
XX
XX
XX Claim 1; SEQ ID NO 692; 1012bp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
XX CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
XX CC ABZ12066) or their mature protein coding portion, active domain coding
XX CC protein or complementary sequences. The polynucleotides are useful for
XX CC identifying expressed genes or for physical mapping of human genome. The
XX CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
XX CC markers, as a food supplement, for generating antibodies, in medical
XX CC imaging, screening and diagnostic assays and for treating cell-
XX CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
XX CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
XX CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
XX CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
XX CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
XX CC arthritis, etc. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,34e-122 Length: 2084
 Score: 2850.00 Matches: 528
 Percent Similarity: 96.4% Conservative: 1
 Best Local Similarity: 96.2% Mismatches: 2
 Query Match: 95.5% Indels: 18
 DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x AB211810 (1-2084)

```

QY      1 MetAapValValGluValAlaGlySerTrpTrpAlaGlnGluAargGluAapTrpIleIleMet 20
DB      163 ATGACGCTGTAGAGAGTGCAGGCGCAGTGTGGGCAACAAGCGAGACGACATCATTTATG 222
QY      21 LysTrpGluValGlyValHisArgAlaGlyLeuProGluAapLysGlyProLysProPheArg 40
DB      223 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAGAGGGGCTTAAGCTTTTCTGA 282
QY      41 SerTrpAenAenAenValAapHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB      283 AGCTACACAAACAACGTGATCATTTGGGANTTGTACATGAGACCGAGCTGCTCTCTG 342
QY      61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAap 80
DB      343 ACTGCGCGGAGCGGACCAATTCGGCGGAGATCAGCCGAAAGACAAAGTGGGTGAT 402
QY      81 MetLeuGlyAapTrpGluLysTrpLysSerSerArgLysLeuIleAapArgAlaTrpLys 100
DB      403 ATGCTGGAGAGACTGGGAGAAATATCAAAAGCACAGCAAAAGCTCATATGATGAGCGTACAG 462
QY      101 GlyMetProMetAenIleArgGlyProMetTrpSerValLeuLeuAenIleGluGluMet 120
DB      463 GGAATGCCCAATGAACATCCGGGGCCGATGTGTGTCAGTCTCTCTGAACATCGAGGAAATG 522
QY      121 LysLeuLysAenProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB      523 AAGTTG----- 528
QY      141 HisIleGlnArgGluLeuAapArgValSerGlyThrLeuAargLysHisIlePhePheArg 160
DB      529 AACATCCAGCGCATCGACCGGAGACGTAAAGCGGACATTAAGAGACATATATTTCTTCAGG 588
QY      161 AapArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
DB      589 GATGATACGGAACCAACAGCAGCGGGAATCATCTCCACATCTCTCTGCAATAGAGAGAT 648
QY      181 AenProGluValGlyTrpCysArgAapLeuSerHisIleAlaIleuPheLeuLeuTrp 200
DB      649 AACCCGGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 708
QY      201 LeuProGluGluAapAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluAargHisSer 220
DB      709 CTTCCTGAGAGAGATGCAATCTCGGCACTGGTGCAGCTGTGGCCAGAGGAGGACATCTCC 768
QY      221 LeuGlnGlyPheHisSerProAenGlyGlyThrValGlnGlyLeuGlnAapGlnGlnGlu 240
DB      769 CTGACGAGGATTTCAACGCCAAATGCGGAGCCGTCCAGGGGCTTCAAGACCAACAGAG 828
QY      241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAapLysLysAapLeuCys 260
DB      829 CATGTGTAGCCACGTCAACACCCAAACCAATGGGGGCATCGAGACAAAGAAAGATCTATGT 888
QY      261 GlnGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAapGlyIleSerLeuGly 280
DB      889 GGGAGAGTGTCCCGTTAGGCTGTCTCATCGGAAATTAATGAGGGGATCTCTCTCGGG 948
QY      281 LeuThrLeuAargLeuTrpAapValTrpLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
DB      949 CTCACCTCGGCCCTGTGGAGCTGTATCTGGTAAGAGCGCAACGCGCTTATGCGCGATA 1008
QY      301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgCysGlyPro 320

```

```

DB      1009 ACAAGAAATCGCCTTTAAAGTTTCAGACAGAGCGCTCTCAAGAAAGCTCCAGGTGTGCCCG 1068
QY      321 TrpAlaArgPheCysAenArgPheValAapThrTrpAlaArgAapGluAapThrValLeu 340
DB      1069 TGGGACAGCTTTTTCACACCGGTTGTGTGATACCTGGGSCCAGGATAGAGACATGTGTCTC 1128
QY      341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAapLeuProProPro 360
DB      1129 AAGCATCTTAGGGCCTCTATGAGAAACTAAACAAGAAAGAGGGGAGCTTGCAACCCCA 1188
QY      361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB      1189 GCCAAACCCGACAAAGGATGTGTGGCATCCAGAGCTGTGGCTTCAAGTGGCGGAGAG 1248
QY      381 ThrLeuCysLysGlyAapArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB      1249 ACCCTCTGCAAGGGGACAGGCAAGGCCCTCTCAAGGCCCAACGCCGGTTCGCCGGGCC 1308
QY      401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
DB      1309 ATTTGTCACTTCCCGCCACCGGACCTGTCTTCCACACCTGTCTCTGTGGTGGGCT 1368
QY      421 ValArgGluAapTrpTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB      1369 GTCCGGAGAACACACTACCTCTGTGGGCACTCAGGGTGTGCCAGCCCGCGCTGGCTCAG 1428
QY      441 GlnGlyProGlnGlySerTrpArgPheLeuGlnTrpAenSerMetProArgLeuProThr 460
DB      1429 GAGAGACTCTAGGGTCTCTGAGATTCCTCAAGTGAATCTCATGCCGCCCTCCCAAG 1488
QY      461 AapLeuAapValGlnGlyProTrpPheArgHisTrpAapPheArgGlnSerCysTrpVal 480
DB      1489 GACCTGACGTAGAGGGCCCTGTGGTCCGCATATATTTCAAGACAGAGCTGCTGGTCTC 1548
QY      481 ArgAlaIleSerGlnGluAapGlnLeuAlaProCysTrpGlnAlaGlnIleProAlaGlu 500
DB      1549 CGTGCATATTCCTCAGAGAGACCAAGCTGCCCCCTGTGGGAGGCTGAACACCTGGCGAG 1608
QY      501 ArgValArgSerAlaPheAlaAlaProSerThrAapSerAapGlnGlyThrProPheArg 520
DB      1609 CGGATGAGATCGGCTTTCGCTGCACCCAGCATGATTCGACACAGGGGACACCCCTTCAGA 1668
QY      521 AlaArgAapGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB      1669 GCTAGGAGCAACACGCGTGTGCTCCACCTCAAGGGGCTTGCTGCGCTTCACCTTG 1728
QY      541 GluSerSerGlnPheProProGlyPhe 549
DB      1729 GAAAGTTCTCAAGTTCCCTCCAGGCTTC 1755

```

RESULT 13
 ADM44328
 ID ADM44328 standard; cDNA; 2084 BP.
 XX AC
 XX ADM44328;
 AC
 XX 03-JUN-2004 (first entry)
 DT
 XX
 DB Novel human arginine-rich protein cDNA #692.
 XX
 XX ser; gene; human; arginine-rich protein; cancer; inflammation;
 KW genetic disorder.
 XX
 XX Homo sapiens.
 XX OS
 XX US2004053250-A1.
 EN
 XX 18-MAR-2004.
 PD
 XX 21-NOV-2002; 2002US-00302172.
 PE
 XX 05-MAR-2001; 2001US-00799451.
 FR 05-MAR-2002; 2002WO-US005095.
 PR

PR 20-AUG-2002; 2002US-00225251.
XX (TANG/) TANG Y T.
PA (XUEA/) XUE A. A.
PA (DRMA/) DRMANAC R T.
XX
XX Tang YT, Xue A, Drmanac RT;
XX WPI; 2004-238579/22.
DR
XX
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
XX
XX Disclosure; SEQ ID NO 692; 51pp; English.
XX
XX The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
XX
XX Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.34e-122 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.4% Conservative: 1
Best Local Similarity: 96.2% Mismatches: 2
Query Match: 95.5% Indels: 18
Gaps: 1
US-10-071-838-2 (1-549) x ADM44328 (1-2084)
QY 1 MetAspValValGluValAlaGlySerTTPAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGAGCGCTGTAGAGGTGGGGGCGATTGGTGGGCAAGAGCGAGGACATCATATATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAGAAGGACACCGAGCTGGGCTCCAGAGACAAAGGGCTTAAGCTTTTTCGA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACAAACAACGTCATCATTTGGGATTTGTATCATGACGAGCGTGGCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTTPValAsp 80
Db 343 ACTGCGGGGAGCGAAGCAAAATTCGGCGGAGATCAGCCGAAGAACAAAGTGGGTGAT 402
QY 81 MetLeuGlyAspTTPGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATCTGGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCATAGTCAGAGCTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCGATGTGGTCAGTCTCTCGAACAACGTAGGAATG 522
QY 121 LysLeuLysAsnProGluArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG-----528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGAAAGCATATATTTCTTAG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 589 GATCGATACGGAACCAAGACGCGGAACTACTCCACATCTCTCGGATATGAGAGTAT 648

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
Db 649 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCACATTCGCCCTTGTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCTGAGGAGATGATTCATTCGGGACTGGTCACTGCTGGCCATGAGAGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTCAGAGGATTTCAACAGCCCAATAGCGGGAGCCGTCCAGGGGCTCCAAACCAAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTATCCACGTCACCAACCAAGACCATGGGATGAGACAAAGAAATCTATGT 888
QY 261 GlyIleCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGGTTCCTCCGTTAGGCTCCCTCATTCGATATTGATGACGGATCTCTCTGGG 948
QY 281 LeuThrLeuArgLeuTTPAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 949 CTCACCTGGGCTGTGGAGCGTGTATCTGTGTAAGGCGAACAGGGGTTGATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATCGCTTAAAGTTCAGAGAACCGCTCACGAAGACGTCAGAGTGGCCG 1068
QY 321 TTPAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGACAGTTTTCGCAACCGGTTCTGTGATACCTGGGCCAGGAGATGAGACATGTCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGCGCTCTATGAGAAACTAAACAAGAGAGGGGGGCTCGAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1189 GCCAAACCCGAGAAAGGTGTGTGGATCCAGGCGCTGTGCGGCTTCACTGCGGGAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGluProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGAGACAGGAGGCCCTCCAGGCCACACCGGTTCCCGGGCC 1308
QY 401 IleTTPSerAlaSerProProAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1309 ATTTGGTCAGCTTCCCGGACAGGGACCTCGTTCTTCCACACCTGTCTGGTGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACCTTACCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGGGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTTPArgPheLeuGlnTTPAsSerPheProArgLeuProThr 460
Db 1429 GAGAGACTCAGGCTTCTGAGATTCCTGCAATGAACTCCATGCGCCGCTCCCAAG 1488
QY 461 AspLeuAspValGlnGlyProTTPPheArgHisTyrAspPheArgLysSerCysTTPVal 480
Db 1489 GACCTGAGATGAGAGGGCCCTTGGTCCCGCATTAATGATTTCAGACAGAGCTGCTGGT 1548
QY 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTTPGlnAlaGlnHisPProAlaGlu 500
Db 1549 CGTGCATATCCAGAGAGACCGAGCGGCCCTCTGTGGAGGCTGAAACCTTGGGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CGGGTAAATCGGCTTTCCTGTCACCCAGCATGATTCGACCAAGGAGCAACCCCTTCAAG 1668
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1669 GCTAGGAGCAAGACAGCGTGTGCTCCACCTGAGGGCCTTGGCGGCTCCCACTTG 1728

QY 401 IletpserAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
 DB 1889 ATTGGTCAGCTTCCCGCCACGCGCACCTGCTTCCACACCTGCTGCGGAGCT 1948
 QY 421 ValArgGluAptThrTyProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 DB 1949 GTCGCGGAAACACCTACCTGCTGGGACATCAGGGTGTGCCACGCGCGCTGCTGAG 2008
 QY 441 GylGlyProGlnGlySerThrPargPheLeuGlnTrpPheSerMetProArgLeuProThr 460
 DB 2009 GAGAGACCTCAGGGTCTTCGAGATTCCTGACAGGAATCCACAGCCCGCTCCCAACG 2068
 QY 461 AspLeuAptValGluGlyProTrpPheArgHisTrpAptPheArgGlnSerCysTrpVal 480
 DB 2069 GACCTGACCTGAGGGCCCTTGTGCTCCGCTATTAATTTCAACACAGAGCTGCTGAGTC 2128
 QY 481 ArgAlaIleSerGlnGluAptGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
 DB 2129 CGTGCATATCCCGAGGAGACACAGCTGCGCCCTGCTGGGACGCTGAACACCTGGGAG 2188
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrArgSerArgGlnGlyThrProPheArg 520
 DB 2189 CGGGTGAGATCGGCTTTCGCTGACCCAGCATGATTCGACAGCGGACCGCTTCAGA 2248
 QY 521 AlaArgAptGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
 DB 2249 GCTAGGACCAACAGCCGCTGTCTCCACCTCCAGAGGCTTGTGCTGCGCTTCACCTTG 2308
 QY 541 GluSerSerGlnPheProGlyPhe 549
 DB 2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335

RESULT 15

AB211809
 ID AB211809 standard; cDNA; 2072 BP.
 XX
 AC AB211809;
 XX
 DT 20-JAN-2003 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 691.
 XX
 XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KM cell-proliferative disorder; neurodegenerative disease; bacterial;
 KM Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KM arthritis; cystostatic; immunomodulator; neotropic; neuroprotective;
 KM antiParkinsonian; antidiabetic; immunosuppressive; dermatological;
 KM haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KM antiarthritic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002MO-US005095.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Dmanac RT;
 XX
 XX WPI: 2002-759812/82.
 DR P-Psdb; ABP69592.
 XX
 XX New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet

PT or coagulation disorders.
 XX
 PS Claim 1; SEQ ID NO 691; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (AB211119-
 CC AB212066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP68902-ABP69843) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, wound, burns, incision, ulcers, liver
 CC platelet or coagulation disorders, viral, fungal, parasitic),
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WFO at ftp.wipo.int/pub/published_pcc_sequences
 XX
 SQ Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2 65e-121	Length:	2072
Score:	2827.00	Matches:	525
Percent Similarity:	95.6%	Conservative:	0
Best Local Similarity:	95.6%	Mismatches:	2
Query Match:	94.8%	Indels:	22
DB:	6	Gaps:	1

US-10-071-838-2 (1-549) x AB211809 (1-2072)

QY 1 MetAspValValGluValAlaGlySerTrpPalaGlnGluArgGluAptIleIleMet 20
 DB 163 ATGACCTGGTAGAGCGCGGCGGACGTTGGTGGCAACAGAGGAGGACATCATTTATG 222
 QY 21 LysTrpGlnLysGlyHisArgAlaGlyLeuProGluAptLysGlyProLysProPheArg 40
 DB 223 AATATGAAAGGAGGACCGAGCTGGCTGCCAGAGCAAGAGGCGCTTATGCTTTTGA 282
 QY 41 SerTrpAspAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
 DB 283 AGCTTACCAACAACCTGATCTTTGGGATTTGTCATGAGCGAGCTGCTCTCTG 342
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
 DB 343 ACTGCCGGGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 402
 QY 81 MetLeuGlyAptTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTrpLys 100
 DB 403 ATGCTGGAGACTGGGAGAAATCAAAAGCAGCAGAAAGCTCATATCGAGCGTCAAG 462
 QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
 DB 463 GGAATGCCCATGAAACATCGGGGCCGATGTGTCAGTCTCTGAACTGAGAAATG 522
 QY 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 DB 523 AAGTTGAAAAACCCGGGAAGATACAGATCATGAAGAGAGGCAAGAGTATCTGAG 582
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 DB 583 CACATCCACGCGCATTCACCGGGGCTGTACCGGGGACCTTAAAGAAATATATCTTCAGG 642
 QY 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
 DB 643 GATCGATACGAAACCAAGCAGCGGAACTACTCTCAATCTCTGCAATGAGAGTAT 702
 QY 181 AsnProGluValGlyTrpCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTrp 200
 DB 703 AACCCGAGGTGGGCTACTGACGAGGACTGAGCAGATGCGCGCTTCTCTCTAT 762

Search completed: April 5, 2006, 14:49:21
Job time : 1222 secs

```
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 763 CTTCTGTGAGAGAGATGCATTCGTGGCACTGGTGCAGCTGCTGGCCAGAGAGGACCTCC 822
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 823 CTGCAGGAGATTTCACAGCCCAATGCGCGGACCGTCCAGGGGCTCCAAAGACCAACAGAG 882
Qy 241 HisValValAlaThrSerGlnProLeuThrMetGlyHisGlnAspGlyLeuAspLeuGly 260
Db 883 CATGTGTAGCAGCTCAACACCCAAAGACCATGGGGCATCAG----- 924
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCTCGGG 936
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 937 CTCAACCTGCGCCCTGTGGGACGTGTATCTGTAGAGAGCGCAACAGCGCTGTATGCCGATA 996
Qy 301 ThrArgIleAlaPheLeuValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACNAGAAATCGCCTTTAAGGTTTCAGCAAGAGCGCTTCACGAAAGACSTCCAGGTGTGGCCG 1056
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1057 TGGGACAGTTTTTGGCAACGGGTGGTGTATACCTGGGCCAGGGATGAGGACACTGTGCTC 1116
Qy 341 LysHisLeuAlaArgLaserMetLysLeuThrArgGlyGlnGlyAspLeuProProPro 360
Db 1117 AAGCATCTTAAAGGCTCTATATGAAACTAACAAAGAAAGCAGGGGACCTGCACACCCCA 1176
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1177 GCCAAACCCGAGCAAGAGGTCGTCCGCAATCCAGGCTGTGCGGCTTCACTGGCGGGAG 1236
Qy 381 ThrLeuCybLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTGTGCAAGGGGGACAGGACAGGCCCTCCAGGCCACCAAGCCGGTCCCGGGGCC 1296
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1297 ATTGTGTCAGCTTCCCGCCACAGGGCACCTCGTTCCTTCACACCTGTCTGTGGGGCT 1356
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1357 GTCCGGAGAGACACTACCTGTGGGCACTCAGGGTGTGCCAGGCCCGGCTGCTCAG 1416
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1417 GAGAGACTCAGGGGTCTTGAGGATTCCTGCAGTGAATCTCATGCCCCGCTCCAAAG 1476
Qy 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1477 GACCTGGAGCTAGAGGGCCCTGTGGTTCGCGCATATGATTTCCAGACAGAGCTGCTGGGTC 1536
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 1537 CGTGCCATATCCCAAGAGAGACCACTGGCCCCCTGTGCAAGGCTGAACACCCCTGGAG 1596
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1597 CGGGTGAATCGGGCTTTGGCTGCACCCAGCACTGATTCGACCAAGGGCACCCCTTCAGA 1656
Qy 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1657 GCTAGGGAGCAACAGCCCTGTGCTCCACTCAGGGGCTGTGCTCTGCGGCTCCACTTGG 1716
Qy 541 GluSerSerGlnPheProGlyPhe 549
Db 1717 GAAAGTTCAGTTCCTCCAGGCTTC 1743
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:42:06 ; Search time 6702 Seconds

(without alignments)
4656.379 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWMAQREDIIM.....TSGPCLCGLHLESGPPPGF 549

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1p

-Q=/abs/ABSSWB.spool/US10071838/runat_05042006_141913_5146/app_query.fasta_1

-DB=GenEmbl -QFMT=fastap -SUFFIX=xge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45

-LOCAL=200 -THR SCORE=pcpr -THR MAX=100 -ALIGN=15 -MODE=LOCAL

-OUTFMT=pcpr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs05p

-USRR=US10071838 -CCGN_1_1_7415 -runat_05042006_141913_5146 -NCPu=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_da:*

2: gb_in:*

3: gb_env:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_sts:*

11: gb_sy:*

12: gb_un:*

13: gb_vl:*

14: gb_htg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2983	100.0	1964	6	AX775945	AX775945 Sequence
2	2983	100.0	1964	8	AF540953	AF540953 Homo sapi
3	2975	99.7	1964	6	AX775947	AX775947 Sequence

4	2971	99.6	2090	8	BC071680	BC071680 Homo sapi
5	2963	99.3	2072	6	AX086847	AX086847 Sequence
6	2945	98.7	2166	8	BC075809	BC075809 Homo sapi
7	2906.5	97.4	2304	6	AX775943	AX775943 Sequence
8	2898.5	97.2	2304	6	AX775941	AX775941 Sequence
9	2850	95.5	2084	6	AR578516	AR578516 Sequence
10	2849	95.5	2575	8	BC033670	BC033670 Homo sapi
11	2835	95.0	2647	6	AX775951	AX775951 Sequence
12	2827	94.8	2072	6	AR578515	AR578515 Sequence
13	2827	94.8	2647	6	AX775949	AX775949 Sequence
14	2827	94.8	2647	6	AX833552	AX833552 Sequence
15	2827	94.8	2647	6	AK095385	AK095385 Homo sapi
16	2732	91.6	2111	8	BC058890	BC058890 Homo sapi
17	2626.5	88.0	1924	8	BC078140	BC078140 Homo sapi
18	2595	87.0	3389	8	AK122833	AK122833 Homo sapi
19	2191.5	73.5	7878	6	AX335512	AX335512 Sequence
20	2191.5	73.5	7878	8	HSTR210	X63546 H. sapiens m
21	2188.5	73.4	7971	8	AY143550	AY143550 Homo sapi
22	2184.5	73.2	8491	8	HSM807865	BX647719 Homo sapi
23	2133.5	71.5	8201	6	I76205	I76205 Sequence 9
24	2133.5	71.5	8201	6	AX774956	AX774956 Sequence
25	2133.5	71.5	8201	8	HSTR213	X63547 H. sapiens m
26	1522	51.0	134506	14	AC146482	AC146482 Homo sapi
27	1522	51.0	180876	8	AC021317	AC021317 Homo sapi
28	1522	51.0	181517	8	AC126327	AC126327 Homo sapi
29	1521	51.0	193646	8	AC131056	AC131056 Homo sapi
30	1518	50.9	146243	14	AC130293	AC130293 Homo sapi
31	1514	50.8	375757	8	AC139091	AC139091 Homo sapi
32	1505	50.5	173053	14	AC129893	AC129893 Homo sapi
33	1501	50.3	140152	8	AC027821	AC027821 Homo sapi
34	1501	50.3	140152	8	AC027821	AC027821 Homo sapi
35	1500.5	50.3	158698	14	AC036181	AC036181 Homo sapi
36	1500.5	50.3	192881	8	AC067923	AC067923 Homo sapi
37	1500	50.3	206040	14	AC132814	AC132814 Homo sapi
38	1497.5	50.2	185449	14	AC139361	AC139361 Homo sapi
39	1494.5	50.1	42073	14	AC144840	AC144840 Homo sapi
40	1493.5	50.1	41266	14	AC139176	AC139176 Homo sapi
41	1493.5	50.1	89247	14	AC146857	AC146857 Homo sapi
42	1487.5	49.9	146243	14	AC130293	AC130293 Homo sapi
43	1469	49.9	123925	8	AC003976	AC003976 Homo sapi
44	1465	49.1	172770	14	AC023133	AC023133 Homo sapi
45	1448.5	48.6	163628	14	AC139367	AC139367 Homo sapi

ALIGNMENTS

RESULT 1	AX775945	1964 bp	MRNA	linear	PAT 14-JUL-2003
LOCUS	AX775945	Sequence 215 from Patent WO03048202.			
DEFINITION	AX775945				
ACCESSION	AX775945.1	GI:32693663			
VERSION	AX775945.1				
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Hominidae; Homo.				
REFERENCE					
AUTHORS	Matsuda, A. and Muramatsu, S.				
TITLE	NP-kb activating gene				
JOURNAL	Patent: WO 03048202-A 215 12-JUN-2003;				
	Aashi Kabei Kabushiki Kaisha (JP)				
FEATURES					
source					
	1..1964				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	1..1650				
	/note="unlabeled protein product"				
	/codon_start=1				
	/protein_id="CAE11672.1"				
	/db_xref="GI:32693664"				

REFERENCE Homidae; Homo.
1 (bases 1 to 1964)
AUTHORS Pei L., Peng Y., Ling X.B., van Eynhoven, W.G., Nguyen, K.C.Q.,
Rubin, M., Hoey, T., Powers, S. and Li, J.
TITLE PRC17, a novel oncogene encoding a Rab GTPase-activating protein,
is amplified in prostate cancer
JOURNAL Cancer Res. (2002) In press
REFERENCE 2 (bases 1 to 1964)
AUTHORS Li, J.L., Peng, Y. and Powers, S.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2002) Genomics, Tularek, Inc, 266 E Pulaeki Road,
Suite 1, Greenlawn, NY 11740, USA
FEATURES
source
1. 1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
1. 1964
/gene="PRC17"
1. 1650
/gene="PRC17"
/product="Rab GTPase-activating protein PRC17"
/protein_id="AA033117.1"
/db_xref="GI:23452655"
/translation="MDVVEVAGSWMAQREDITMKYEKGRAGLPEDKPKPRSTNN
NVDHLIVHETLPLEPLAREAKQIRREIISKRWMDLGMWETKYSRKIDAYKCM
PNIRGPMWSVLNIEBMLKNGROYIMKEKRSSEHIQIDRVSGLRGIIFR
DRYGTQKRELHILAVYEYNPVCGCRDLISHIALFLYLPEEDAFMALVQLASER
HSLQGFSPKSGTVQGLDQDQEHVATSOPTMHODKXDLGGCGSPGLGLIILIDG
ISGLTLRLMDVLYSGEOLMPITRIAFVQOKRLTKSRCCPMARFCRPFDTYAR
DEDTVTLKRLASMKKLRKQDLPPEPKSQSSASRVPASRGKTLCKGSDQAPFG
PPARFPRPIWSAPRAPRSTPCPGAVVEDTYPVGTQVPSPALAQGGQDSMFL
QNMVSMPLPTDLVEGFWPRHYDFRQSCWRAISOEDQLAPCWQAEBHPRVSAFPA
PSTSDQGTFRARDEPCAPTSGLCLGLHSSQPPPGF"

ORIGIN

Alignment Scores:
Pred. No.: 9,43e-95 Length: 1964
Score: 2993.00 Matches: 549
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-071-838-2 (1-549) x AF540953 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTTPAlaGlnGluArgLysArgLysSerSerGlu 20
DB 1 ATGGACGTGGTAAAGGTCGGCGGACATGGTGGGCACAAAGCGAGGACATCATTTATG 60
QY 21 LysTyrGlnLysGlnHisArgAlaGlyLeuProGlnLysArgLysProLysProPheArg 40
DB 61 AAATTCGAAAGGAGGACCGACGCTGGGCTCCAGAGGACAAAGGGCCTTAAGCTTTTCCA 120
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGlnLeuProProLeu 60
DB 121 AGCTAACAAACAACGTCGATCTTTGGGATTTGTACATAGACGAGGAGCTCCCTCTG 180
QY 61 ThrAlaArgGlnLysGlnIleArgArgGlnLysSerArgLysSerLysTTPValAsp 80
DB 181 ACTGCCGGGAGCGGAAGAAATTCGGCGGAGATCGCCGAAAGACAAAGTGGGTGGAT 240
QY 81 MetLeuGlyAspTTPGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 241 ATCTCGGAGACCTGGGAGAAATACAAAGACAGCAAGAAAGCTCATAGTCAAGCTACAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGlnGluMet 120
DB 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTCTGAAACATTTGAGAAATG 360

QY 121 LysLeuLysAsnProGlnArgTyrGlnIleMetLysGlnLysArgLysSerSerGlu 140
DB 361 AAGTTGAAAACCCCGGAAGATACAGATCATTAAGAGAAAGGCAAGGCTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 421 CATATCCAGCGCATACACCGGACGTAAGCGGACATTAAGAAAGCATATATCTTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgLysLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
DB 481 GATCGATACCGAACCAAGACGCGGAATCATCTCAATCTCTTGGCATATGAGGAGTAT 540
QY 181 AsnProGlnValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
DB 541 AACCCGAGGTGGGCTACTGACGAGACCTGAGCCATCCGCCCTTTGCTCTCTAT 600
QY 201 LeuProGlnLysAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 601 CTTCCTGAGGAGATCATCTGGGACCTGTGACACTGCTGGCCAGTGAGAGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
DB 661 CTGCAGAGATTTCACAGCCCAATGGCGGACCGTCACAGGCGCTCCAAAGCAACAGAG 720
QY 241 HisValValAlaThrSerGlnProLysTTPMetGlyHisGlnAspLysAspLeuCys 260
DB 721 CATGTGTACCCACGTCACAAACCAACCATGGGCGATCGAGACAAAGAAATCTATGT 780
QY 261 GlyIleCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 781 GGGCAGATGTTCCCGTTAAGGCTCCTCATCCGATATTGATTGACGAGGATCTCTCGGG 840
QY 281 LeuThrLeuArgLeuTTPAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 841 CTCACCTCGCGCTGGGACGCTGATCTGTGAAGGCGMAAGCGCTTGAGCGCGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgCysGlyPro 320
DB 901 ACAAGATGCGCTTTAAGTTACGACAAAGCGCTTACAAAGCGTCCAGGTGTGGCCG 960
QY 321 TTPAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspLysAspThrValLeu 340
DB 961 TGGGCACTGTTTGGCAACCGGTCGTTGATACCTGGGCGAGGATGAGGACATGAGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnLysAspLeuProProPro 360
DB 1021 AAGCATCTTAAGGCTCTTATGAAGAACTTAACAAAGAGGAGGACCTGCCACCCCA 1080
QY 361 AlaLysProGlnGlnLysSerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1081 GCCAAACCCGAGCAAGGTCGTCGATCCAGGCTGTGCGGCTTCACTGCGGGAAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1141 ACCCTTGCAAGGGGAGACAGGAGGCTCCACAGGCCCAACAGCCGGTTCGGGCGCC 1200
QY 401 IleTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
DB 1201 ATTTGTCAGCTTCCCGGACAGGCACTCGTCTTCCACACCTGCTCTGTGTGGGGCT 1260
QY 421 ValArgGlnAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCGGGAAAGAACCTTACCTGTGGCACTCAAGGAGTGTGCCAGCCGGGCTGTGGCTCAG 1320
QY 441 GlyIleProGlnGlySerTTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
DB 1321 GAGAGACCTCAGGCTTCTGAGATTCCTGAGGTGAATCCATGCCCCGCCCTCCAAAG 1380
QY 461 AspLeuAspValGlnGlyProTTPPheArgHisTyrAspPheArgLysCysTTPVal 480
DB 1381 GACCTGAGAGTGAAGGGCCCTTGGTTCGCCCATTAATTAATTCAGACAGGCTGTGGGTC 1440
QY 481 ArgAlaIleSerGlnLysAspGlnLeuAlaProCysTTPGlnAlaGlnIleProAlaGln 500

Db 1441 CGTGCATATCCACGAGAGACAGCTGGCCCCCTCTGGCAGGCTGAACACCTTGGGAG 1500

Qy 501 ArgValaTgSerAlaPheAlaAlaProSerThrApsSerApsGlnGlyThrProPheArg 520

Db 1501 CGGGTGAATCGGCTTGTGCTGCACCCAGCACTGATTCGACCAAGGAGACCCCTTTCGA 1560

Qy 521 AlaArgApsGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540

Db 1561 GCTAGGAGCAAGACCGCTGTGCTCCACCTCAGGCGCTTGTGCTGTGCGGCTCCACTTG 1620

Qy 541 GluSerSerGlnPheProGlyPhe 549

Db 1621 GAAAGTCTCAGTTCCTCCAGGCTTC 1647

RESULT 3

AX775947 1964 bp mRNA linear PAT 14-JUL-2003

LOCUS Sequence 217 from Patent WO03048202.

DEFINITION AX775947

ACCESSION AX775947

VERSION AX775947.1 GI:32693665

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Matveuda, A. and Muramatsu, S.

AUTHORS NF-KB activating gene

TITLE Patent: WO 03048202-A 217 12-JUN-2003;

JOURNAL Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

location/Qualifiers

source

1. 1964

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1. 1650

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE11673.1"

/db_xref="GI:32693666"

/translation="MDVVEVAGSMWAQERED1IMKYEKHRAGLPEBKGPFRSYNN

NVDHIGIVHETELPPLTARBAQIRRETSRSKSKWMDMGDEKYSRKLIDRAYKGM

PMNIRGPMWSVLTINIEWKLNKGRVQYIMKEKGRSSHHIORIDYSGTRKHIFPR

DRYCTKQRELIHIIILAYERYVBYCYCDLSHIALPLTYIPREDAFWALVQLASR

HSLOGFSPNGSTVQGLDQDEHVAVISQPTMGHODKDDCGGCSPLGLIRLLDS

ISLGLTLRLMDVYLVEGQALMPTTRIAFKVQKRLTTTSRCGFWARCNRFVDTMAR

DEDIVLKLIRASMKKLTRKQGLQPPAPAEQSSASRPVPSRGKTLCKGDRQAPRG

PPAPFPRIMGASPPRARSSTPCFGGAVREDTYPVGTQGVSPALAQGGQSMRFL

QWNSMPRLPTDLVGEPMWRHYDFRQSCWVARISQEDQLAPCQWAEHAEVRSAFPA

PSTSDQCTPRRADDEPCATPSRCGLHLESSQPPRG"

ORIGIN

Alignment Scores:

Pred. No.: 1.78e-94 Length: 1964

Score: 2975.00 Matches: 548

Percent Similarity: 99.88 Conservative: 0

Best Local Similarity: 99.88 Mismatches: 1

Query Match: 99.74 Indels: 0

Gaps: 0

US-10-071-838-2 (1-549) x AX775947 (1-1964)

Qy 1 MetApsValValGlnValAlaGlySerTrpTrpAlaGlnGlnValArgGlnApsIleIleMet 20

Db 1 ATGACACGTGTAGAGGTGCGGCGAGTGTGCGGCAAGAGCGAGAGACATCATATATG 60

Qy 21 LysTrpGlnValGlnHisArgAlaGlyLeuProGlnApsIleGlyProCysProPheArg 40

Db 61 AAAATACGAAAGGAGCACCGAGCTGGCTGCCAGACGAAAGGGGCTTAACCTTTTCGA 120

Qy 41 SerTrpApsApsValApsHisLeuGlyIleValHisGlnTrpGlnLeuProProLeu 60

Db 121 AGTACAAACAAACGTCGATCATTTGGGATTTGATCATGACGAGGCTGCTCTCTG 180

Qy 61 ThrAlaArgGlnAlaValGlnIleArgArgGlnIleSerArgIysSerLysTrpValAps 80

Db 181 ACTGCGGAGAGGGAAGCAAAATTCGGCGGAGATCCGGAAGACCAAGAGTGGTGGAT 240

Qy 81 MetLeuGlnApsTrpGlnLysTrpLysSerSerArgIleApsArgAlaTrpLys 100

Db 241 ATGCTGGAGATCTGGAGAAATACAAAGCAGCAAGAAAGCTCATATGATCGAGGCTACAG 300

Qy 101 GlyMetProMetCaspIleArgGlyProMetTrpSerValLeuLeuApsIleGlnIleMet 120

Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCAGTCTCTCCGAACATTTGAGGAATG 360

Qy 121 LysLeuLysApsProGlnArgTrpGlnIleMetLysGlnLysValApsSerSerGln 140

Db 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGCAAGAGCTCATTTGAG 420

Qy 141 HisIleGlnArgIleApsArgApsValSerGlyThrLeuArgLysHisIlePhePheArg 160

Db 421 CACATCCAGGCGCATCGACCCGGAGCTAAGCGGACATTAAGAAAGCATATATCTTCAGG 480

Qy 161 ApsArgTrpGlyThrLysGlnArgGlyLeuLeuHisIleLeuLeuAlaTrpGlnGlnTrp 180

Db 481 GATCATACGGAACCAAGCAGCGGGAACCTACTCCATCTCTGCGCATATGAGAGATAT 540

Qy 181 ApsProGlnValGlyTrpArgApsLeuSerHisIleAlaAlaLeuPheLeuLeuTrp 200

Db 541 AACCCGAGAGTGGCTACTGACAGGACCTAAGCACATCCCGCTTGTCCTCTTAT 600

Qy 201 LeuProGlnGlnApsAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnApsHisSer 220

Db 601 CTTCCTAGAGAGATGATGATCTGGGCACTGTGTCAGCTGTGCGACGTGAGAGGCACTCC 660

Qy 221 LeuGlnGlyPheHisSerProApsArgGlyThrValGlnGlyLeuGlnApsGlnGlnGln 240

Db 661 CTGACGAGGATTTTCACAGCCCAAAATGGCGGAGCCGCTCCAGGCGCTCCAAAGCAAGAG 720

Qy 241 HisValValAlaThrSerGlnProLysThrMetGlnHisIleGlnApsLysValApsLeuCys 260

Db 721 CATGTGGTACCCAGCTACAAACCAAGACATGGGAGATAGAGCAAGAAAGATCTATGT 780

Qy 261 GlnGlnCysSerProLeuGlnGlyCysLeuIleArgIleLeuIleApsGlyIleSerLeuGly 280

Db 781 GGGAGAGTTCCTCCGTTAAGCTGCTCATCCGATATTTATGACGGGACTCTCTCTGGG 840

Qy 281 LeuThrLeuArgLeuTrpApsValTrpLeuValGlnGlyGlnAlaLeuMetProIle 300

Db 841 CTCACCTGCGGCTGTGGACGTGTATCTGTAGAAAGGCGAAACAGGCGTTGATGCCGATA 900

Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

Db 901 ACAAGAAATGCCCTTTAAGTTTCAGCAGAAAGCGCTCCAGAAAGCTCCAGATGGGCGCG 960

Qy 321 TrpAlaArgPheCysApsApsPheValApsThrTrpAlaArgApsGlnApsTrpValLeu 340

Db 961 TGGGACAGTTTTCGAAACCGGTTCTGTGATACCTGGGCGGAGATAGAGACATGTGCTC 1020

Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyApsLeuProProPro 360

Db 1021 AAGCATCTTAGGGGCTCTATGAAGAACTAACAAGAAAGCAGGGGCGACTGCAACCCCA 1080

Qy 361 AlaLysProGlnGlnGlnLysSerSerAlaSerArgProValProAlaSerArgGlyLys 380

Db 1081 GCCAAACCCAGCAAGAGGTGTGCGCATCCAGGCTGTGCGGCTTCACTGTGGGGAGAG 1140

Qy 381 ThrLeuCysLysGlnApsArgGlnAlaProProGlyProProAlaArgPheProArgPro 400

Db 1141 ACCCTCTGCAAGGGGAGCAGGAGGCTCTCCAGGCCACCAAGCCGCGTTCGCCGGGCC 1200

Qy 401 IleTrpSerAlaSerProArgAlaProArgSerSerThrProCysProGlyGlyAla 420

DB 1201 ATTGGTCAGCTTCCCGCCAGCGGACCTGTTCTTCAACACCTGTCTCGGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCCGGGAAACACCTACCTGTGGGACATCAGGGTGTGCGACGCCGGCTGGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTPArGpHeuGlnTTPAnSerMetProArGLeuProThr 460
DB 1321 GGAGGACCTCAGGGTCTCTGAGATTCCTCACTGAGATTCATCCGCCCTCCCAAG 1380
QY 461 AspLeuAspValGluGlyProTTPpHeaArgHisTyrAspPheArgGlnSerCyStrVal 480
DB 1381 GACCTGGACCTGAGGGGCTTGGTTCGCCATTATGATTTCAAGACAGACTGTGGTTC 1440
QY 481 ArgAlaHisSerGlnGluAspGlnLeuAlaProCyStrTPGlnAlaGlnHisProAlaGlu 500
DB 1441 CGGCAATATCCAGGAGGACGAGCTGCCCCCTGCTGGAGGCTGAAACCCCTGCGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1501 CGGTGATGATCGGCTTTCCTGCAACCGACATGATTCGACACAGGACCCCTTCAGA 1560
QY 521 AlaArgAspGluGlnProCyValAProThrSerGlyProCyLeuCyGlyLeuHisLeu 540
DB 1561 GCTAGGAGACGAAACAGCGTGTGCTCCCACTCAAGGGCTTGGCTTGGGCTTCACCTTG 1620
QY 541 GluSerSerGlnPheProPheGlyPhe 549
DB 1621 GAAAGTTCATGATTCCTCCAGGCTTC 1647

RESULT 4
LOCUS BC071680 2090 bp mRNA linear PRI 08-MAR-2005
DEFINITION Homo sapiens TBC1 domain family, member 3, mRNA (CDNA clone
ACCESSION MGCL87891 IMAGE:5165385), complete cds.
VERSION BC071680
KEYWORDS BC071680.1 GI:47939522
SOURCE MGC.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 2090)
Strasberg RU, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
B, Bueltow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Heish F, Dietchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P,
Pirange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaby
SJ, Bosnak KH, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richard S, Worley KC, Hale S, Garcia AM, Gay LJ, Hultik SM,
Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Helton E, Kettman M, Madan A, Rodrigues S, Sanchez A, Whitting M,
Maden A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,
Touchman DW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
Schmütz J, Myers RM, Butlerfield JS, Krzywinski MI, Skalska U,
Smalutz DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2090)
Director MGC Project.
Direct Submission
Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk

Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehg.stanford.edu>
Contact: (Dickson, Mark) medpaxi1.stanford.edu
Dickson, M., Schmütz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 166 Row: g Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984.
Location/Qualifiers
1. 2090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87891 IMAGE:5165385"
/cissue_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 2090
/gene="TBC1D3"
/note="synonyms: PR17, TRE17"
/db_xref="GeneID:84218"
/db_xref="MIM:607741"
101. 1750
/gene="TBC1D3"
/codon_start=1
/product="TBC1 domain family, member 3"
/protein_id="AAH71680.1"
/db_xref="GI:47939523"
/db_xref="GeneID:84218"
/db_xref="MIM:607741"
/translation="MDVVEASGMAOEREDITIMKYEKGRAGLPEDKPKPRSYNN
NVNDCGIVHETELPPLTAREAKQIREISRKSKVMDLDMWEYKSRKLDIAVGM
PNKINGPMWSVLINTEMLKNDGRQIKERKRSSEHRIORDRDISGRLDMFPR
DRYCKORRLHILAYREYNPYCGRLSHALFILYLPEDAFMLVOLLASER
HSIQGHSFNGCTVQGLDQOQHVAATSPKTKMGHDKDLCQCGSPGLCLTITLIDG
ISLGLTLRLMDVLYLVEGEQALMPITRIARKVQKRLTKTSRCGPMARFCNRFVDTWAR
DBDTPVLRASMKLTKRQGLPPAPKEQSSASRPVPSRGKTLCKGRQAPPG
PPARFPRPIWASPPRAPRSSTPCGAVREDTYPVGTGVSPPALACGPGOSMPL
QNSMPLRPLTDIVSGPMFRHYDFRGSQVRAISORDLAPCQAEHPARVRSAPFA
PSTDSQCTPPPARBDQCAFPSGRLCLGLHSSQPPGP"

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.:	2,6e-94	Length:	2090
Score:	2971.00	Matches:	546
Percent Similarity:	99.8%	Conservative:	2
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	8	Gaps:	0

US-10-071-838-2 (1-549) x BC071680 (1-2090)

QY 1 MetAspValAlaGluValAlaGlySerTPTPAlaGlnGluArgGluAspIleIleMet 20
DB 101 ATGACGTGTGATGAGCTCGCGGCACTTGTGGGCAAGAGAGGAGACATCATTAATG 160
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysPheArg 40
DB 161 AAATACGAAGAGGACACGAGCTGGCTGCCAGAGACAGGAGGCTTAAGCTTTTCGA 220
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGlnThrGluLeuProLeu 60
DB 221 AGCTACACACACAGCTCATCTTTGGGATTTGTATCATGAGACGAGCTGCTCTTG 280

QY	61	ThrAlaArgGluAlaIleGlnIleLeuArgArgGluIleSerArgIleSerIleThrValLeu	80
Db	281	ACTGGCGCGAGGCGCAAGCAAAATTCGGCGGGAGATCACTCCGAAAGCAAGCGGTGAT	340
QY	81	MetIleuGlyAspTrpGluIleYerTrpYssSerSerArgIleLeuIleAspArgAlaTrpYlys	100
Db	341	ATGCTGGGAGACTGGGAGAAATACAAAGCAGCAAAAGCTCATAGATCGAGCGTACAG	400
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlyMet	120
Db	401	GGAAATGCCCATGAAATCATCCGGGGCCCGATGTGTGATGATCTCTGAAATGAGAAATG	460
QY	121	LysIleuLysAsnProGluYArgTrpGlnIleMetLysGluIleYsGlyLysArgSerGlu	140
Db	461	AAAGTGAAGAAACCCCGAAGATACAGATCATGAAGAGAGAAAGGAGGTCATCTAG	520
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
Db	521	CACATCCAGCGCATGACCCGGGACATTAAGCGGGACATTAAGAGACATATGTTCTTCAG	580
QY	161	AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluTrp	180
Db	581	GATGATACGGAAACCAAGACGGGGAATCATCTCAATCTCTCGCATATAGAGATAT	640
QY	181	AsnProGluValGlyTrpCysArgArgPheSerHisIleAlaIlePheLeuLeuTrp	200
Db	641	AAACCGAGGTGGGTCTACTGCAGGAGACCTGACCCATCTCGCGTTGTCTCTCAT	700
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluIleArgHisSer	220
Db	701	CTTCTGAGAGAGATGCATTCATTCGGGCACTGGTGCAGCTGTCTGGCAGTGAAGGCACTCC	760
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	761	CTGCGAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGCTCCAAAGCAACACAGAG	820
QY	241	HisValValAlaIleThrSerGlnProIleYsrThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	821	CATGTGGTAGCCACGTCAACACCAAGACCAAGGGGACATCAGGACAAAGAAAGTCTTAAT	880
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	881	GGGCACTGTTCCCTTATGAGCTGCTCTCATCCGATATGATTGACGGGATCTCTTCGGG	940
QY	281	LeuThrLeuArgLeuTrpAspValTrpLeuValGlnGlyGlnAlaLeuMetProIle	300
Db	941	CTCACCTCGCGCTGTGGGACGTGATCTGGTAGAAGGCAACAGCGTGTATGCCATA	1000
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	1001	ACAAAGATCGCTTTTAGGTTCAGCAGAAAGCCCTCAAGAAAGCTTCAGAGGTGGCCCG	1060
QY	321	TrpAlaArgPheCysAsnArgPheValAspTrpTrpAlaArgAspGluAspThrValLeu	340
Db	1061	TGGGACGTTTTTGCACCGGTTGTTGATACCTGGGCAAGGATAGAGCACTGTGCTC	1120
QY	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1121	AAAGCATCTTAAAGGCTCTTATGAAGAACTAACAAGAAAGCAGGGGACCTGCACCCCA	1180
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys	380
Db	1181	GCCAAACCCGAGCAAGGGGTGCTGGCATCCAGGCTGTGGCGGCTTCACTGGCGGGAG	1240
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1241	ACCCTCTCAAGGGGAGCAGGAGGCGCTCTCAGGCCCAACAGCCGGTTCCTCCGGGCCC	1300
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysAspGlyGlyVala	420
Db	1301	ATTGTGTACGCTTCCCGCGCAGGGGCACTGCTTTCTTCACACCTGTGCTGTGTGGGCT	1360
QY	421	ValArgGluAspThrTrpTrpValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440

Df	1361	GTCCGGGAGACACTTACCCTTGCGGCACTCAGGGATGTGCCAAGCCCGGCTGGCTANG	14220
OY	441	GLYGIProGInglYSerTPRPaRPhelengIntPranserMetProArgLeuProThr	460
Df	1421	GGAGGACCTCAGGGGTTCTCGAGATTCCTCGCAGTGAATCATTCATGCCCCGCTCCAAAG	1480
OY	461	AapleuApVaLlGlUgLYPProTPRPhearghIstYrhapPheargGlnserCyatTPVaL	480
Df	1481	GACCTGCAGCGTAGAGGGCCCTTGATTCGCGCATTAATGATTTCACAAGAGGTGCTGGGGTC	1540
OY	481	ArGaLaalieserGIngluaSPGLLeuAlAProSerThraPseAerpGInglYthrProPhearg	500
Df	1541	CCTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGGCTGAACACCTTCGGAG	1600
OY	501	ArgValARgsereAlaphaelaalAProSerThraPseAerpGInglYthrProPhearg	520
Df	1601	CGGGGAGATCGGGCTTTGGCTGCACCCAGCACTGATTCGACAGGGGACCCCTTCAGA	1660
OY	521	AlaarharpGInglInProCYaALAProThiserGlyProCYaleuCYeGLyleuHileu	540
Df	1661	GCTAGGGGCGAAACAGCAAGTGTGCTCCACCTCAGAGGCGCTTGCTCTGGGCTCCACTTG	1720
OY	541	GluserSerGlnPheProPGlyPhe	549
Df	1721	GAAGTTCTCAGTTCTCTCCAGGCTTC	1747
RESULT 5			
AX086847		2072 bp	DNA linear PAT 09-MAR-2001
LOCUS	AX086847	Sequence 799 from Patent WO0112659.	
DEFINITION	AX086847		
ACCESSION	AX086847.1	GI:13276047	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	Wiemann, S.		
AUTHORS	Human dna sequences		
TITLE	Patent: WO 0112659-A 799 22-FEB-2001;		
JOURNAL	German Human Genome Project (DE)		
FEATURES	Location/Qualifiers		
source	1..2072		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	4,86e-94	Length:	2072
Score:	2963.00	Matches:	545
Percent Similarity:	99.6%	Conservative:	2
Best Local Similarity:	99.3%	Mismatches:	2
Query Match:	99.3%	Indels:	0
DB:	6	Gaps:	0
US-10-071-838-2 (1-549) x AX086847 (1-2072)			
OY	1	MetaApVaLVAlGUuVAlAgLYSeRTPTPLagngluarGluAepILleleMec	20
Df	99	ATGACGCGTAGAGGTGCGCGGCGCATGTTGGCGCACAAAGCGAGGACATCATTAATG	158
OY	21	LysrTYrGluYusGlyhIsARGAlAGLYLeuProGluAepLySGlyProLYsProPhaRg	40
Df	159	AAAATACGAAAGGACACCGAGCTGGGTGCACAGGACAAAGGGGCTTAAAGCTTTTCGA	218
OY	41	SetTYrAsnaAnanVanAlaSpHisLeugLYlleValHAGluThrgLUleuProbleu	60
Df	219	AGCTACAAACAAACGTCGATCATTTGGGGATTGATCATGACGAGAGCTGCTCTCTG	278

Qy	61	Thrlaargjuaalalyglnlleargrglulleaserarglyserlytrpvalasp	80
Db	279	ACTCGGGGAGGAGGAGCAAAATTCGGGGAGATCAGCCGAAAGAGAAATGGGTGAT	338
Qy	81	MetleuglyabpTrrpGlylyserlyserarglyleuulleapargalalylys	100
Db	339	ATGCTGGAGAGCTGGGAGAAATACAAAGACAGAGAAAGCTCATAGATCGAGCTACAG	398
Qy	101	GlywerPrometAenlleargglyPrometTrrserValleuleuAenlleuglUmet	120
Db	399	GGAATGCCCATGAAATCCGGGGCCCATGTGTGATCTCTGAAACACTGAGAAATG	458
Qy	121	LyseleuylabpProglYargTrrglnllemetlysglylylyargserglu	140
Db	459	AAGTGAAGAAACCCCGAAGATACCAATCATGAAAGAAAGGCAAGAAATCATTCAG	518
Qy	141	Hisllleglnarglleapargaspvalserglythreuaarglyshilephepharg	160
Db	519	CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAAAGCATATATCTTCAG	578
Qy	161	AsparglyrrglythrrlyrglnargluleuuleuhsilleleuuleuallatrrglnlUtr	180
Db	579	GATCGATACCGAACCAAGACGGGAACTACTCCACATCTCTGGCATATAGAGAGTAC	638
Qy	181	AspProgluValglyTrrCyargAspPleuserhsillealaleuPheleuLeuTyr	200
Db	639	AAACCGAGGTGGGCTACTCGAGGAGCTGAGCCATCCGCGCTGTCTCTCTAT	698
Qy	201	LeupProgluValaapalaphetrrpaleuValgleuleuualasergluarghseser	220
Db	699	CTTCCTGAGAGGATGATTCGTGGGACCTGGTGCAGCTGTGGCAGTGAAGGACATCC	758
Qy	221	LeuglnllyphenhseserProargglylythrrValglnllyleuglnaarglnlglu	240
Db	759	CTGAGGAGATTCACACCCCAAAATGGCGGACCTCCACAGGGGCTCCCAAGACCAAGAG	818
Qy	241	HisValValaThrSerGlnProlyeThrMetGlyhsllaglnaaplyblyAspLeuCy	260
Db	819	CATGTGTAGCCAGTCAACCAACCAAGCCATGGGGCATCGAGCAAGAAAGATCTATGT	878
Qy	261	GlyGlnCyaserProleuglyCyaleuullearglleuulleapglylleserleugly	280
Db	879	GGGACAGTGTCCCGTTAGGCTGGCTCATCGGATATTGATTGACGGGATCTCTCGGG	938
Qy	281	LeuThrleuargleuTrrpaspvalTrrleuValgluglylaleuMetProile	300
Db	939	CTCACCTCGCGCTGTGGGACGTGTATCTGTAGAAAGCGGAACGGGCTGATGCCGATA	998
Qy	301	ThrArgllealaphelyValglnllyleuArgleuThrrlyserThrserargCyarglyPro	320
Db	999	ACAAAGATCGCTTTAAAGTTACGACAAAGGCGCTCAAGAAAGCTCAGGTGTGGCCG	1058
Qy	321	TrrpalaargPheCyasaenAargPheValaspTrrTrrpalaargaspgluaspTrrValleu	340
Db	1059	TGGGCAAGTTTTCGAAACCGGTTCTGTGATACCTGGGCGAGGATGAGACATGTCTC	1118
Qy	341	LyehlsleuarglaserMetlylyleuThrrarglysglnlyaspLeuProPro	360
Db	1119	AACATCTTATGGGCTCTATGAAAGAACTAAACAAGAAAGGGGGGACCTGCCACCCCA	1178
Qy	361	AlalyerProgluGlnlyserSerAlaserArgProalProalaserargglylylyl	380
Db	1179	GCCAAACCCGAGGAGGTGTGCGATCCAGGCTGTGCGGCTTACAGTGGGGGAG	1238
Qy	381	ThrluCyalyseglYasparGlnalaProProglYProProalAargPheProargPro	400
Db	1239	ACCTCTGCAAGGGGAGGAGGAGGCGCTCCAGGCGCCACAGCCGCTTCCGCGGCC	1298
Qy	401	IlletrrserAlaserProProarglaProargserThrProCyasProglYlyl	420
Db	1299	ATTGTGACACTTCCCGGACAGGCGACCTGTCTTCCACACCTCTCTGTGGGGCT	1358
Qy	421	ValArggluaspThrrTrrProValglnllyThrglnllyValProserProalaleuAlaGln	440

Db	1359	GTCCGGAGAGCACTTACCTGTGGGACATCAGGCTGTGCCACCCGCGCTGCTAG	1418
Qy	441	GlyGlyProGlnlyserTrrpargPheleuglnlTrrpanserMetProargLeuProThr	460
Db	1419	GAGAGACCTCAGGGGTCTCTGAGATCTCTGAGATGGAATCCATGACCCCGCTCCAAAG	1478
Qy	461	AspLeuaspValgluglyProTrrpPhearghsierlytrpPhearglninserCyatrrVal	480
Db	1479	GACCTGACGTAGGGGCGCTGTGGTCCGCAATTATATTTCAACAGAGCTGTGGGTC	1538
Qy	481	ArgAlalleserGlnlyAspArglnleuAlaProCystrrrpAlaGlnhsProalaglu	500
Db	1539	CGTCCATATCCAGAGAGACAGACTGGCCCCCTGTGGGAGCTGAACACCTTCGGAG	1598
Qy	501	ArgValargSerAlaPheAlaAlaProSerThrAspSerAspGlnlyThrProPhearg	520
Db	1599	CGGGTGAATCGGCTTTCGTGCACCCAGACATATTCGACAGGAGCAACCCCTTCAGA	1658
Qy	521	AlaArgaspGlnlgnProCyasAlaProThrSerGlyProCyaleuCyarglyleuhsleu	540
Db	1659	GCTAGGAGACAAACAGAGTGTGTCTCCACCTCAGGGGCTGTGCTCGGCTCCACTTG	1718
Qy	541	GlyserSerGlnPheProProglYpHe	549
Db	1719	GAAAGTTCTCAGTTCCCTCCAGGCTTC	1745
RESULT 6			
BC075809			
LOCUS			
DEFINITION			
Homo sapiens TBCL domain family, member 3B, mRNA (cDNA clone			
MGC:87892 IMAGE:5170064), complete cds.			
BC075809			
MGC:87809.1 GI:49902676			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Homnidae; Homo.			
1 (bases 1 to 2166)			
Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G.,			
Klauner, R.D., Collins, F.S., Wagner, L., Shlemen, C.M., Schuler, G.D.,			
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhac, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.J., Wang, J., Hsieh, F.,			
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
Scheltz, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S.,			
Carninci, P., Prange, C., Raha, S.S., Loguettano, N.A., Peters, G.J.,			
Ardamson, R.D., Mullaly, S.J., Bosak, S.A., McEwen, P.J.,			
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Foley, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,			
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Bouffard, G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Gilmour, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smalins, D.E.,			
Schnerch, A., Schein, J.E., Jones, S.J., Jones, S.J., and Marra, M.A.			
Mammalian Gene Collection Program Team			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
12477932			
2 (bases 1 to 2166)			
NITH MGC Project			
Direct Submission			
Submitted (06-JUL-2004) National Institutes of Health, Mammalian			
Gene Collection (MGC), Bethesda, MD 20892-2590, USA			
NTH-MGC Project URL: http://mgc.ncl.nih.gov			
Contact: MGC help desk			
Email: cgabbs-remail.nih.gov			

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at: <http://image.lnu.gov>
Series: IRAX Plate: 166 Row: h Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 1414984.

FEATURES

SOURCE

Location/Qualifiers

gene

CDS

1..2166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87892 IMAGE:517064"
/tissue_type="Brain, adult medulla"
/clone_id="N1H_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2166
/gene="TRCID3B"
/note="synonym: PRC17"
/db_xref="GeneID:414059"
147..1796
/gene="TRCID3B"
/codon_start=1
/product="TRC1 domain family, member 3B"
/protein_id="AAH75809.1"
/db_xref="GI:49902677"
/db_xref="GeneID:414059"
/translation="MDVVEVAGSWMAQEREDIMKYEKGRAGLPEDKGPFRSYNN
NVDHGIYHETELPPLTARBAKQIRREISRKWDMDGMEKYSRKLIDRAYKGM
PMNIRGPMWSTVLNIEBKLNKGRVIOIKMEKGRSSHRIORIDISGLRBMFPR
DRYCTKORELHILILAYEVNPEVGYCDLSHIALPLIYPERDAFVALVOLASRR
HSLOGFSPNGSTVQGLDQOEHVATISQSTMGHODKDCGCCPGLCITRLIDR
ISLGLTLMDVLYVEGQALMPITRIAFAKQORLITKTSKCGWARKCNFVITWAK
DDDTVLKHLRASMKLTKQGLDLPAPAEQSSASRPVPSRGKTLCKGDRQAPPG
PPAPPRPIMSASPPRARSSTPCGSGAREDTYPVGTGVSPALAGGSGMRF
QWNSMPRLPTDLVGEWFRHYDFRQSGAVRAISOEDLAPCQWAEHRAEVRAPAA
PRTSDGCTPRRADBOGCARTSPCLGLHESQPPPG"

ORIGIN

Alignment Scores:

Pred. No.: 2,116-93 Length: 2166
Score: 2945.00 Matches: 541
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.5% Mismatches: 4
Query Match: 98.7% Indels: 0
Gaps: 0

US-10-071-838-2 (1-549) x BC075809 (1-2166)

QY 1 MetAapValValAGluValAlaGlySerTTPAlaGlnGluValGluAapPLeIleMet 20
Db ATGACCGTGTAGAGGTGCGCGGTAGTGTGGGCAAGAGCGAGAGACATCATTTATG 206
QY 21 LyeTYrGluValGlyValGlyValGlyValGlyValGlyValGlyValGlyValGlyVal 40
Db AAATACGAAAGGACACCGAGCTGGCTGCCAGAGGAGCAAGGGGCTTAAGCTTTTGA 266
QY 41 SerTYrAaenAaenValAaPHisleuGlylleValHisGluThrGluLeuProleu 60
Db AGCTAACAAACAACGTCATTTGGGATTTGACATGAGACGAGACCTGCTCTG 326
QY 61 ThrAlaargGluAlaValGlnIleArgArgGluIleSerArgValSerLeuTYrValAap 80
Db ACTGCGCGGAGGCGAAGCAATTCGCGGAGATCAAGCGAAGAGCAAGTGGGTGAT 386

QY 81 MetLeuGlyAaPTrpGluValTYrLysSerSerArgLysLeuIleAaPArgAlaTYrLys 100
Db ATGCTGGAGAGACTCGAGAAATACAAAGACAGCAAGAAAGTCATGATTCAGGCTACAG 446
QY 101 GlyMetProMetAaenIleArgGlyProMetTTPSerValLeuAaenIleGluGlnMet 120
Db GGAATGCCATGAAACATCCGGGGCCGATGGTCTCTCCGAAACATTGAGGAATG 506
QY 121 LysLeuValAaenProGluValArgTYrGlnIleMetLysGluValArgSerSerGlu 140
Db AAGTTGAAAAACCCGAAAGATACACATCATGAAAGAAAGGCAAGAGTCTTCAG 566
QY 141 HisIleGlnArgIleAaPArgAaPValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db CACATTCAGGCGAATCGACCGGACATTAAGGGGACATTAAGAGGACATATGTTCTCAG 626
QY 161 AaPArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db GATGATACGGAACCAAGACAGCGGAATCTACATCTCTCGCATATGAGAGCAT 686
QY 181 AaPProGluValGlyTYrCysArgAaPLeuSerHisIleAlaAlaLeuPheLeuTYr 200
Db AACCCGAGGTGGCTACTCGAGGACCTAGCCATCGCCCTTGTTCTCTCTAT 746
QY 201 LeuProGluGluAaPAlaPheTrpAlaLeuValGlnLeuLeuAaSerGluArgHisSer 220
Db CTTCCTAAGAGAGATGATCTTGGGCACTGTGGAGTGGTGGCACTGAGAGCATCTCC 806
QY 221 LeuGlnGlyPheHisSerProAaenGlyTYrValGlnGlyLeuGlnAaPArgGlnGlu 240
Db CTGACAGGATTCACACCCCAATGGCGGAGCCCTCCAGGGGCTCCAAAGCAACAGAG 866
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAaPLeuValAaPLeuCys 260
Db CATGTGTAGCACGTACATCCAAAGCATGGGGCATGAGCAAGAAAGATCTATGT 926
QY 261 GLyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAaPArgIleSerLeuGly 280
Db GGGAGAGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTGACCGGATCTCTCGGG 986
QY 281 LeuThrLeuArgLeuTYrAaPValTYrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db CTCACTCTGGCGCTGTGGAGCTGTATCTGTAGAAAGCAAGCGGCTGATGCCATTA 1046
QY 301 ThrArgIleAlaAaPLeuValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db ACAAGAAATCGCTTTAAGTTCAAGACAGAGCGCTCCAGAAAGCTCAGGTGTGGCCA 1106
QY 321 TrpAlaArgPheCysAaenArgPheValAaPTrpAlaArgAaPArgValLeu 340
Db TGGGCAAGTTTTCGAAACCGGTTGTTGATACCTGGGCAAGGATGATGACACTGTGCTC 1166
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAaPLeuProPro 360
Db AAGCATCTTAGGCGCTTATGAAAGAACTAAAGAAAGCAAGAGGCGGACCTGCCACCCA 1226
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProAlaProAlaSerArgGlyLys 380
Db GCCAAACCCGAGCAAGGTGTGCGCATCCAGGCTGTGCGGCTTACGTTGCGGGAGAG 1286
QY 1227 GCCAAACCCGAGCAAGGTGTGCGCATCCAGGCTGTGCGGCTTACGTTGCGGGAGAG 1286
Db ACCCTTCGCAAGGGGACAGGAGCGCCCTTCAGAGCCCAACAGCCGGTTCGCGGGCCC 1346
QY 401 IleTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
Db ATTGTGACAGTTCCTCCGCAACCGGCACTGTTCTTCCACACCTCTGCTGGTGGGCT 1406
QY 421 ValArgGluAaPTrpTYrProValGlyThrGlnGlyValProSerProAlaLeuValGln 440
Db GTCCGGGAAAGACACTTCCCTGTGGGCACTCAGGGTGTGCCAGCCGCGCTGTGCTAG 1466

```

Qy 441 G1G1YpROgInG1YSeRtPArGpHeuGInTrPAsSeRMeCProArGLeuPProThr 460
Db 1467 GGAAGAGCTTCAGAGTTCCTGGAGATTCTCGACATGGAATCCATGCCGCCCTCCCAAG 1526
Qy 461 AAPLeuAPValG1G1YpROtRPhEaRgH1eTyRAsPHeaRgInSeRyGtRPA1 480
Db 1527 GACCTGGACGTAGAGGGGCTTGCTCCGCCATTATGATTGACAGAGAGCTGCTGGGTC 1586
Qy 481 ATGAlAl1eSeRgInG1uAPrG1nLeuAlAPrOCyTtRg1nAlAG1uH1SPrOAlAG1u 500
Db 1587 CGTGCCATATCCCAAGAGAGACAGCTGGCCCCCTGTGGCAGGCTGAACACCTCGCGAG 1646
Qy 501 ATGValARgSeRAlaPHeAlaAlAPrOCeRThrAsPSeRAsPqInG1YThRProPHeArG 520
Db 1647 CGGGTAgATCGGCTTTCCTGTCACCCGGCAGCTGATTCGACACAGGGCAGCCCTTCACA 1706
Qy 521 AlARgAPrG1uGInPrOCyAlAPrOThrSeRg1YpROCyAlEucYsG1YLeuH1AlEuc 540
Db 1707 GCTAGGAGCAAGACAGAGTGTGCTCCACCTCAGGGCCTTGCTGCGGCTCCACTTG 1766
Qy 541 G1uSeRSeRgInPHeProPProG1YpHe 549
Db 1767 GAAAGTTCAGATTCCCTCCAGGCTTTC 1793

RESULT 7
LOCUS AX775943 2304 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 213 from Patent WO03048202.
ACCESSION AX775943
VERSION AX775943.1 GI:32693661
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 213 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source 1..2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB11671.1"
/db_xref="GI:32693662"
/translation="MDVVEVAGSMWAQEREDIIIMKEYKGRAGLBEDKPKPRSYNN
NVDHIGIVHETELPPLTARBAKQIRREISRSKWDMLGDMKYSRKLIDRAYKM
PMNIGPMWVSVELNTEEMKLNKGRYQJIMKEGKSSSEIQRIDDSGTLAKHIEFR
DRYGTGKRELIHILAYERYNPEVATGCRDLSHIALFLIYLEEDAFAWLYLLASBR
HSLGDFHSPNGSTVGLDDODDEHVAVATGPOKTMGHODKDDCGGSSPIGCLRILIDG
ISLGTITLMDVYLYVGEALMPTIRIAFKVQOKRLTKTSRGPWAFRCNRPVWAW
DEDVTLKLRASMKLTKRGQDLPPAPPEQSSASRPVPARGKTLCKGRQAPG
PFGLPRHGHVLPHPVLYVGLSGKPTLMALVCPARPWLREDLRPVDSGCTPCPA
SORTWT"

ORIGIN
Alignment Scores:
Pred. No.: 4,766-92 Length: 2304
Score: 2906.50 Matches: 540
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 2
Query Match: 97.4% Indels: 8
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AX775943 (1-2304)
Qy 1 MetAPValValG1uValAG1YSeRtPTrPA1AGInG1uArgG1uAP1le1leMet 20

```

```

Db 362 ATGAGAGGTGATAGAGTCCCGGCACTGTGTGGCCCAAGAGGAGACATCATTAAG 421
Qy 21 LyeTYrG1uLyeG1YH1sARgAlAG1YLeuProG1uAPuYsG1YProLYsProPHeArG 40
Db 422 AATACGAAGAAAGGACACCGAGCTGGCTGCCAGAGGACAAAGGGGCTTAAGCTTTTCCA 481
Qy 41 SeTYrAsnAsnAsnValAsnH1sLeuG1Y11eValH1sG1uThR1uLeuProPLeu 60
Db 482 AGCTTACAAACAAACGTCATCATTTGGGATTTGATCAATGAGCGAGCTGCTCTCG 541
Qy 61 ThrAlARg1uAlAlYsG1u11eARgAG1u11eSeARg1YSeRtYsTPValAsP 80
Db 542 ACTGCGCGGAGGCGCAAGCAATTCGCGCGAGATACCGCGAAAGCAAGTGGTGGAT 601
Qy 81 MetLeuG1YAsPTP1uLYsTYrLYsSeRSeRArgLYsLeu11eAPARgAlATYLYs 100
Db 602 ATGCTGGAGAGACTGGAGAAATATCAAAACACAGAAAGCTCATAGATCAGGCTACAG 661
Qy 101 G1YMeTPromETAsn11eARg1YPromETTPSeRVal1LeuLeuAsn11eG1u1uMet 120
Db 662 GGAATGCCATGAAACATCGGGGCCGATGTGGTCAGTCTCTGAACACTGAGGAAATG 721
Qy 121 LyeLeuLYsAsnProG1YARgTYrG1u11eMetLYsG1uLYsG1YLYsARgSeRSeRg1u 140
Db 722 AAGTTGAAAACCCCGAAGATACCGATCATGAAGAGAAAGGCAAGGTCATCTGAG 781
Qy 141 H1s11eG1nARg11eAPARgAPValSeRg1YThRLeuARgLYsH1s11ePHeARg 160
Db 782 CACATCAGCGCATGACCGGAGCGTAAACGCGGAGCATTAAGAAAGCATATATCTTCAGG 841
Qy 161 AsPARgTYrG1YThRLYsG1nARg11uLeuLeuH1s11eLeuLeuAlATYrG1uG1uTYr 180
Db 842 GATCATTAAGAAACCAAGAGCGGAACTATCCATCTCTCTGCAATATGAGAGATAT 901
Qy 181 AsnProG1uValAG1YTYrCYsARgAsPLeuSeRHS11eAlAlAlEucPHeuLeuTYr 200
Db 902 AACCGGAGGTGGGCTACTGCGAGGACCTGAGCCATGCGCGCTTGTCTCTCTCAT 961
Qy 201 LeuProG1uG1uAPARgAPHeTTPAlAlEucValG1nLeuLeuAlAsERG1uARgH1sSeR 220
Db 962 CTTCCTGAGAGAGATGATCTTGAGGCACTGGTGCAGCTGCTGCGCATGAGAGGCACTCC 1021
Qy 221 LeuGInG1YpHeH1sSeRProAsnG1YG1YThRValG1nG1YLeuG1nAsPqInG1u 240
Db 1022 CTGCAAGGATTTCAAGCCCAATGCGGAGCGTCCAGGGGCTCCAAACCAACAGGAG 1081
Qy 241 H1sVal1AlAlAThSeRgInPrOlySThRMeTg1YH1sG1nAsP1YLYsAsPLeuCYs 260
Db 1082 CATGTGTAGCCACGTCACAAACCCAAAGCATGGGGCATCAGAGCAAGAAAGTCTATGT 1141
Qy 261 G1YcInCYsSeRProLeuG1YCYsLeu11eARg11eLeu11eAsPq1Y11eSeRLeuG1Y 280
Db 1142 GGGCAGGTGTTCCCGCTTAGGCTGCTCATCCGGAATATTGACGGGATCTCTCGGG 1201
Qy 281 LeuThRLeuARgLeuTTPAsPValTYrLeuValG1uG1YLeuG1nAlAlEucMetPro11e 300
Db 1202 CTCACCTCGCGCTTGAGACGTGTATCTGTAGAAAGGCAAGAGCGGTGATGCGGATA 1261
Qy 301 ThrARg11eAlAPHeLYsValG1nG1nLYsARgLeuThRLYsThSeARgCYsG1YPro 320
Db 1262 ACAGAAATGCGCTTTAAGGTTTCAGCAAGAGCGCTTCAGAAACGTCAGAGTGTGGCCCG 1321
Qy 321 TPAlARgPHeCYsAsnARgPHeValAsPTThRTPAlARgAsPq1uAsPTThRValLeu 340
Db 1322 TGGGACAGTTTGTGCAACCGGTTCTGTATCTGAGGCGAGGAGTGAAGCACTGTGGCTC 1381
Qy 341 LYsH1sLeuAlARgAlAsESeMeLYsLeuThRArgLYsG1nG1YAsPLeuProProPo 360
Db 1382 AAGCATCTTAAGGCTCTATGAAGAACTTAAGAAAGGAGGAGGAGCACTGCAACCCCA 1441
Qy 361 AlAlYsProG1uG1YSeRSeRAlAsERARgProValProAlAsERARg1YG1YLYs 380

```

Db 1442 GCCAAACCCGAGCAGGAGGTCGCGCATCCAGGCGCTGCGGCTTCACGTCGCGGAG 1501
Qy 381 ThleucUetYveGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1502 ACCCTCTGCAAGGGGAGCAGGACAGCCCTCCAGGC-----CCA 1540
Qy 401 lletPserAlaSerProProArgAlaProArgSerSerThProCyProGlyGlyAla 420
Db 1541 TT-TGGTCACCTTCCCGCCAGCGGCACTGCTTCCACACCCCTGCTCCGAGGCGCT 1599
Qy 421 ValArgGluAphThyTyProValGlyThGlnGlyValProSerProAlaLeuAlaGln 440
Db 1600 GTCCGGGAGACACCTACCTGTCGACCTCAGGGTGTGCCACCCGCGCTGCGTCAG 1659
Qy 441 GlyGlyProGlnGlySerThProArgPheLeuGlnTrpAphSerMetProArgLeuProThr 460
Db 1660 GAGAGACCTCAGGGTCTCTGAGATCTCTGAGTGAATCCAGTCCAGCCCGCTCCACAG 1719
Qy 461 AspLeuAphValGlyGlyProTrpPheArgHisTyAspPheArgGlnSerCySTrpVal 480
Db 1720 GACCTGACCTAGAGGCGCTTGGTTCGCCATTATGATTTCAAGACAGAGCTGTGGTTC 1779
Qy 481 ArgAlaIleSerGlnGluAphGlnLeuAlaProCySTrpGlnAlaGlnHisProAlaGlu 500
Db 1780 CGTGCCTATCTCCAGAGAGACAGCTGCGCCCTGCTGAGGCTGAACACCCCTGCGGAG 1839
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThraPheSerAphGlnGlyThProPheArg 520
Db 1840 CGGATGAGATCGGCTTTCGCTGACCCAGCATGATTCGACACAGGCGACCCCTTCAGA 1899
Qy 521 AlaArgAphGlnGlnProCyAlaAlaProThSerGlyProCyLeuCySGlyLeuHisLeu 540
Db 1900 GCTAGGACCAACAGCGCTGTGCTCCACCTCAGGGCGCTTCCCTGCGGCTTCACCTTG 1959
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 1960 GAAAGTTCAGTTCCTCCAGGCTTC 1986

RESULT 8
AX775941
LOCUS AX775941 2304 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 211 from Patent WO03048202.
ACCESSION AX775941
VERSION AX775941.1 GI:32693659
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 211 12-JUN-2003;
Aashi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. 2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11670.1"
/translation="MDVYVAGSWMAQERDIIIMKYKGRHAGLPEDKGPKEPFSYNN
NVHIGIVHETELPPLTAREAKQIRRIISKSKVMDLGMWETKYSRKLIDRAYKGM
PMNIRGPMWSVLTATEMKLKNPGROYIMKSGKSEHIORIDVSGTLRKHIIFR
DRYCKORBELHILAYEYNPEVGYCRDISHIALALVLPEEDAMVALIOLASER
HSLOGFHSNPGCTYQGIQDOOEHNVAATSPRTMGOHKDKLSCGCSPLGCLIRLLIG
ISLGLTLRLMDVYLVESGQALMPTIRIAFVQQRKLTKTSKGFWARCNRFVDYWAR
DEDTVLKHLASMKKLTTRKQGLDLPKAPKPEGSSASAPVSPASRSGKTLCKGDRAPPG
PFQGLPRHGLVLRHPLYLVGISGKTPTLMLALVCAPRPMWLSEDLRVPDGSCTGPCPA

ORIGIN SORTWT"
Alignment Scores:
Pred. No.: 8.98e-92 Length: 2304
Score: 2898.50 Matches: 539
Percent Similarity: 98.2% Conservative: 0
Best Local Similarity: 98.2% Mismatches: 3
Query Match: 97.2% Indels: 8
DB: Gaps: 1
US-10-071-838-2 (1-549) x AX775941 (1-2304)
Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAphIleIleMet 20
Db 362 ATGAGCGTGTATAGAGGTCCGCGGCACTGGTGGGCAAGAGCGAGGACATCATATG 421
Qy 21 LysTyGlnGlyGlyHisArgAlaGlyLeuProGlnAphGlyProGlyPhePheArg 40
Db 422 AAATACGAAAGGAGACACGAGCTGGGCTGCCAGAGGACAAAGGGGCTTAAGCCTTTTCA 481
Qy 41 SerTyraAphAphAphValAphHisLeuGlyIleValHisGlnThrgGluLeuProProLeu 60
Db 482 AGCTACACAAACACGTCGATCTTGGGATTTGTACATGACGAGACTGCTCTCTG 541
Qy 61 ThrAlaArgGlnAlaGlyGlnIleArgArgGlnIleSerArgLysSerTySTrpValAsp 80
Db 542 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAAGAGCAAGTGGTGGAT 601
Qy 81 MetLeuGlyAphTrpGlnGlyTyTyLysSerSerArgGlyLeuIleAphArgAlaTyTyLys 100
Db 602 ATGCTGGAGACTGGAGAAATACAAAGACGACGAAAGCTCTTAATCGAGGCTACAG 661
Qy 101 GlyMetProMetAphIleArgGlyProMetTrpSerValLeuLeuAphIleGlnGlnMet 120
Db 662 GGAATGCCATGAACATCCGAGGCCGATGTGTCTCTCTTGAACACTGAGAAATG 721
Qy 121 LysLeuLysAphProGlyArgTyGlnIleMetLysGlnGlyLysArgSerSerGln 140
Db 722 AAGTTGAAAACCCCGGAGATACAGATCATGAAAGAGAGGCAAGATCATCTAG 781
Qy 141 HisIleGlnArgGlyLeuAphArgAphValSerGlyThLeuArgGlyHisIlePhePheArg 160
Db 782 CACATCCAGGCAATCGACCGGAGCTAAGGAGGACATTAAGAAACATTAATCTTCAGG 841
Qy 161 AspArgTyGlnGlyThryLysGlnArgGlnLeuLeuHisIleLeuLeuAlaTyGlnGlnTy 180
Db 842 GATCGATACGGAACCAAGACGCGGAACTACCTCACATCTCTTGGCATATGAGAGTAT 901
Qy 181 AsnProGlnValGlyTyTyCysArgAphLeuSerHisIleAlaAlaLeuPheLeuLeuTy 200
Db 902 AACCCGAGGTGGGCTATCTGCAAGGACCTGAGCAATCGCCGCTTGTCTCTCTAT 961
Qy 201 LeuProGlnGluAphAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnValGlnHisSer 220
Db 962 CTTCCTGAGGAGGATGATCTGTGGCACTGGTCACTGTGCGCAGTGAAGAGCACTCC 1021
Qy 221 LeuGlnGlyPheHisSerProAphGlnGlyThValGlnGlnGlyLeuGlnAphGlnGln 240
Db 1022 CTCGAGGATTTTCAAGCCCAAAATGGGGGACGCTCCAGGGGCTCCAAAGCAAGAG 1081
Qy 241 HisValValAlaThSerGlnProLysThMetGlyHisGlnAphLysLysAphLeuCys 260
Db 1082 CATGTGTACCAAGCTACAGCAACCAAGACATGGGGGATCAGAGCAAGAAAGATCTATGT 1141
Qy 261 GlyGlnCysSerProLeuGlyTyLysLeuIleArgIleLeuIleAphGlyIleSerLeuGly 280
Db 1142 GGGCAGAGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTACGGGATCTCTCTGGG 1201
Qy 281 LeuThLeuArgLeuTrpAphValTyTyLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
Db 1202 CTCACCTGCGCTGTGGAGCTGTATCTGTAGAAAGCAAGCGCTTGAATGCCATTA 1261

QY 301 ThrArg11eAlaPheIysValGlnGlnIysArgLeuThrIysThrSerArgCysGlyPro 320
 Db 1262 ACAAGATCCCTTTAAGGTTCAAGACAGCGCTTCAAGAGACTCAAGTGTGAGCCG 1321
 QY 321 TTPAlaArgPheCysAenArgPheValAspThrTTPAlaArgAspGluAspThrValLeu 340
 Db 1322 TGGGCAAGTTTGGCAACCGGTTCTGTGATACCTGGGCGAGGATGAGACACTGTCTC 1381
 QY 341 LysH1sLeuAlaArgIAspMetIlybIysLeuThrArgIySglnGlyAspLeuProPro 360
 Db 1382 AACATCTTAAGGCGCTTATGAAGAACTAACAGAAAGAGGAGGAGCTTGCACACCCCA 1441
 QY 361 AlaIysProGlnGlnIysSerSerAlaSerArgProValProAlaSerArgIyGlyLys 380
 Db 1442 GCCAACCCGACAGAGGAGTGTGGACTCCAGGCTTGTCCGGCTTACAGTGGCGGAG 1501
 QY 381 ThrLeuCylsGlyAspArgGlnAlaProProGlyProProAlaArgPheArgPro 400
 Db 1502 ACCCTCTGCAAGGGGAGCAGGAGGCGCTCCAGGC-----CCA 1540
 QY 401 I1eTTPSerAlaSerProProAlaAlaProArgSerSerThrProCysProGlyGlyAla 420
 Db 1541 TT-TGGTCAGCTTCCCGCCACCGGCACTCGTCTTCCACACCTGTCTCGTGGGGCT 1599
 QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1600 GTCCGGAAAGACCTTACCTGTGGGACTCAGAGGAGTGTGCCAGCCCGGCGCTGGCTCAG 1659
 QY 441 G1yGlyProGlnIySerTPArgPheLeuGlnTTPAsnSerMetProAlaLeuProThr 460
 Db 1660 GAGAGACCTCAGGGGTTCTTGAGATTCTGCAAGAACTCCATGCCCGCCCTCCCAAG 1719
 QY 461 AspLeuAspValGluGlyProTTPPheArgH1sTyrAspPheArgIAsnSerCysTPVal 480
 Db 1720 GACCTGACCTTAAGGGCCCTTGTGTTCCGCAATTAATTCAGACAGACCTCTGGGTC 1779
 QY 481 ArgAl1eSerGlnIyAspGlnIyAspGlnIyAspGlnIyAspGlnIyAspGlnIy 500
 Db 1780 CGTGCATATCCAGAGAGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1839
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnIyThrProPheArg 520
 Db 1840 CGGGTAGATCGGCTTCCGCTGACACCCAGACATTCGACACAGGAGGAGGAGGAGGAG 1899
 QY 521 AlaArgAspGlnIyProCysAlaAlaProThrSerGlyProCysLeuCysGlyLeuH1sLeu 540
 Db 1900 GCTAAGGAGACGAAAGCGGTGTCTCCACCTCAGGGGCTTGGCTTGGGCTCCACTTG 1959
 QY 541 GluSerSerGlnPheProGlyPhe 549
 Db 1960 GAAAGTTCTCAGTTCCTCCAGGCTTC 1986
 RESULT 9
 AR578516 2084 bp DNA linear PART 14-DEC-2004
 LOCUS AR578516
 DEFINITION Sequence 692 from patent US 6783969.
 ACCESSION AR578516
 VERSION AR578516.1 GI:56581312
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 2084)
 AUTHORS Tang, Y.T., Goodrich, R.W., Asundi, V. and Drmanac, R.T.
 TITLE Cathespin V-like polypeptides
 JOURNAL Patent: US 6783969-A 692 31-Aug-2004;
 Nuvelo, Inc.; Sunnyvale, CA
 FEATURES
 source location/Qualifiers
 1..2084
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN

Alignment Scores:
 Pred. No.: 3,83e-90 Length: 2084
 Score: 2850.00 Matches: 528
 Percent Similarity: 96.4% Conservative: 1
 Best Local Similarity: 96.2% Mismatches: 2
 Query Match: 95.5% Indels: 18
 DB: 6 Gaps: 1
 US-10-071-838-2 (1-549) x AR578516 (1-2084)
 QY 1 MetAspValAlaGluValAlaGlySerTTPTPAlaGlnGlnIyAspLeuAspI1eMet 20
 Db 163 ATGAGCTGTGTAAGCTCCGGGCACTGTGGGCAAGAGGAGGAGCATCATTAATG 222
 QY 21 LysTyrGluIySglnH1sArgAlaGlyLeuProGluAspIySglnProIyAspPheArg 40
 Db 223 AAATACGAAAGGAGACACCGAGCTGGCTGCCAGAGGACAAAGGGGCTTAAGCTTTTCA 282
 QY 41 SerTTPAspAsnValAspH1sLeuGlyI1eValH1sGlyThrGlyLeuProProLeu 60
 Db 283 AGCTACACAAACAGCTCATCTTGGGATGTACATGAGACGAGCTGCTCTTG 342
 QY 61 ThrAlaArgGluAlaIySglnI1eArgArgGluI1eSerArgIySerIyTPValAsp 80
 Db 343 ACTGCCCGGAGGCGAAGCAATTCGGGAGATCAGCCGAAGACAGAGTGGTGAT 402
 QY 81 MetLeuGlyAspTPGluIyTyrIySglnIySglnIySglnIySglnIySglnIySgln 100
 Db 403 ATCTGGAGAGCTGGAGAAATACAAAGACAGAGAAAGCTCATGATCAGCGTACAG 462
 QY 101 GlyMetProMetAsnI1eArgGlyProMetTPSerValLeuLeuAsnI1eGluGluMet 120
 Db 463 GGAATCCCATGAACATCCGGGCGCGAGTGTGAGTCTCTCTTGAACACTGAGGAATG 522
 QY 121 LysLeuIyAsnProGlyArgTyrGlnI1eMetIyGluIySglnIySglnIySglnIy 140
 Db 523 AAGTTG----- 528
 QY 141 H1sI1eGlnArgI1eAspArgAspValSerGlyThrLeuArgIyH1sI1ePhePheArg 160
 Db 529 AACATCAGCGCATCGACCGGAGCGTAACGAGGACATTAAGAAAGCATATATCTTCAGG 588
 QY 161 AspArgTyrGlyThrIySglnIySglnIySglnIySglnIySglnIySglnIySgln 180
 Db 589 GATCATAGAGAAACAAAGAGCGGGAATCTTCAATCTCTCTGCAATAGAGAGTAT 648
 QY 181 AsnProGluValGlyTyrCysArgAspLeuSerH1sI1eAlaAlaLeuPheLeuLeuTyr 200
 Db 649 AACCGGAGGTGGGCTACTGCAAGGACCTGAGGACATGCGGCTTGTCTCTCTTAT 708
 QY 201 LeuProGluGluAspAlaPheTPAlaLeuValGlnLeuLeuAlaSerGluArgI1sSer 220
 Db 709 CTTCCTGAGAGAGATCATTTCTGGGCACTGTGGAGCTGTGGCCAGTAGAGGCACTCC 768
 QY 221 LeuGlnIyPheH1sSerProAsnGlyGlyThrValGlnGlyLeuGlnIyAspGlnIy 240
 Db 769 CTGCAGAGATTTCACGCGCAATAGCGGAGCGGTCCAGAGGCTCCAAAGCAAGAG 828
 QY 241 H1sValValAlaThrSerGlnProIySglnIySglnIySglnIySglnIySglnIySgln 260
 Db 829 CATGTGTAGCCACGTCACAAACCCAGACATGGGGCATCAGGACAAAGAAATCTATGT 888
 QY 261 GlyIySglnSerProLeuGlyCysLeuI1eArgI1eLeuI1eAspGlyI1sSerLeuGly 280
 Db 889 GGGCAGATGTTCCCGTTAGGCTGCTCATCCGAGATTTATGACCGGATCTCTCGGG 948
 QY 281 LeuThrLeuArgLeuTPAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
 Db 949 CTCACCTTCGCGCTGTGGAGCGTATCTGTGTAAGAGGAGCAAGCGCTTGAATCGGATA 1008
 QY 301 ThrArgI1eAlaPheIysValGlnGlnIySglnIySglnIySglnIySglnIySglnIy 320
 Db 1009 ACAAGATCGCTTTAAGGTTCAAGAGAGAGGCGCTTACAGAAACGTCAGAGTGTGGCCGG 1068

```

Qy      321  Trrp1aaRphCyAsmArpPheValaAspThrTrrp1aaRphGluAspThrValIleu 340
Db      1069  TGGGCAAGCTTTTTCGACACGGTCTGTTGATACCTGGGCGACAGGATGAGACACTGTGCTC 1128
Qy      341  Lysh1aLeuArya1aSerMetLylYleuThraRylGylGlnGlyAspLeuProProPro 360
Db      1129  AAGCATCTTAAAGGGCTCTATGAGAACATAACAGMAAGGAGGGGAGCTTCAACCCCA 1188
Qy      361  Ala1yPpProGluGlnGlySerSerAlaSerArpProVal1ProAlaSerArpGlyGlyYls 380
Db      1189  GCCAAACCCGAGCAGAGGTCCTGGCATTCAGGCTGTGCGGCTTCACTGGGGGAGAG 1248
Qy      381  ThrLeuCyb1ySg1yAspArpGlnAlaProProGlyProProAlaArpPheProArpPro 400
Db      1249  ACCCTCTGCAAGGGGAGCAGGAGCCCTCCAGGCCACAGCCCGCTTCCCGGGGCTC 1308
Qy      401  11eTrrpSerAlaSerProProArpAlaProArpSerSerThrProCybProGlyGlyYla 420
Db      1309  ATTGGTCAGCTTCCCGCCAGCGGAGCCTGCTTCTTCACACCTGCTGCTGGGGGCT 1368
Qy      421  ValArpGluAspThrTrpProValGlyThrGlnGlyVal1ProSerProAlaLeuAlaGln 440
Db      1369  GTCCGGAGAGACACCTTACCTGTGGGCACTCAGGGGTGCCAGCCCGGCTGCTGAG 1428
Qy      441  Gly1yProGlnGlySerTrpArpPheLeuGlnTrpAsnSerMetProArpLeuProThr 460
Db      1429  GGAAGACCTCAGGGGTCTCTGAGATTCCTGCAAGTGAATCTCAGTCCCGCTCCAGAG 1488
Qy      461  AspLeuAspVal1GluGlyProTrpPheArpGlnTrpAspPheArpGlnSerCybTrpVal 480
Db      1489  GACCTGAGCAGTGAAGGGCTTGTGGTCCGCAATATGATTTCAAGACAGCTGCTGGGCTC 1548
Qy      481  Arg1a1aSerGlnGluAspGlnLeuAlaProCybTrpGlnAlaGln1ProAlaGln 500
Db      1549  CTTGCCATATCCAGAGAGACCAAGCTGAGCCCTGCTGCTGAGGCTGAACACCTGGGAG 1608
Qy      501  ArgVal1aRgSerAlaPheAlaAlaProSerThraPseArpGlnGlyThrProPheArp 520
Db      1609  CGGGTGAATCGGGCTTTTGGCTGACACCCAGCACTGATTCGACAGGAGACCCCTTGCA 1668
Qy      521  AlaArpAspGluGlnProCybAlaProThraSerGlyProCybLeuCybGlyLeuHleu 540
Db      1669  GCTGGGAGCAGAACCGCTGTGCTCCACCTCAGAGGCTTCTGCTGCGGCTCCACTTGG 1728
Qy      541  GluSerSerGlnPheProProGlyPhe 549
Db      1729  GAAAGTTCTCAGTTCCCTCCAGGCTTC 1755

RESULT 10
BC033670      2575 bp      mRNA      linear      PRI 09-MAY-2005
LOCUS      Homo sapiens TRC1 domain family, member 3C, mRNA (cDNA clone
DEFINITION      MGC:44903 IMAGE:5167394), complete cds.
ACCESSION      BC033670
VERSION      BC033670.1 GI:45708750
KEYWORDS      MGC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 2575)
REFERENCE
AUTHORS      Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Holtkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Palmer A.A., Rubin G.M., Hong L.,
Seaplton M., Soares M.B., Bonaldo A.F., Casavant T.L.,
Schaefer T.E., Brownstein M.J., Udell T.B., Tashiro Y.,
Carninci P., Prange C., Kana S.S., Loquellano N.A., Peters G.J.,
Aramason R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,

```

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalton, D.K., Muzny, D.M., Sodegrem, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodriguez, S.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M., Skalski, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., Skalski, U., Smalins, D.E.,
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
12477932
2 (bases 1 to 2575)

NIH MGC Project
Direct Submission
Submitted (02-UTL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgti.nih.gov
Ahter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Khong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Maestrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Startipop, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

FEATURES
source
1. .2575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:44903 IMAGE:5167394"
/tissue_type="Brain, adult medulla"
/clone_id="NIH MGC_119"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .2575
/clone="TBC1D3C"
/notes="synonym: MGC44903"
/db_xref="GeneID:414060"
78..1208
/clone="TBC1D3C"
/codon_start=1
/product="TBC1 domain family member 3C"
/protein_id="AA33670.1"
/db_xref="GI:45708751"
/db_xref="GeneID:414060"
/translaction="MDVETVAGSWAQDERDITMKYKGRACLPEDKGPXPRSYNN
NVHDLIVHETLPLTARPAQRIEISKSKSVWMLGWERYKSKRLIDAVKGM
PNMIRGMSVLTNPILKAKQIRIEISKSKSVWMLGWERYKSKRLIDAVKGM
DRGTQKRELLHLLAEVYEVGYCRDISHAALFLVLPEDAFWALVQLASR
HSIOGHSRPMGTVGLQDOQHEVAVTSKTMGHQDKDLGCGSPGLGLRLIDG
VRHRETLAQSPSLPSALLPQCGIAHSGPTGSGSGSPDQAKSAQSGSGQGNW
GSGPASMADWARKLRQAQSLKPMASPSLQEPHPSNBECP"

ORIGIN

Alignment Scores:

Pred. No.:	5,08e-90	Length:	2575
Score:	2849.00	Matches:	543
Percent Similarity:	77.4%	Conservative:	2
Beet Local Similarity:	77.1%	Mismatches:	4
Query Match:	95.5%	Indels:	155
DB:	8	Gaps:	1

US-10-071-838-2 (1-549) x BC033670 (1-2575)

```

QY      1 MetAspValValGluValAlaGlySerTTPTrpAlaGlnGluuArgGluuAspIleIleMet 20
Db      78 ATGGACCTGGTAGAGGTCCGGGGTAGTTGGTGGGCAAGAGCGAGACATCATATTATG 137

QY      21 LysTyrGluuysGlyHisArgAlaGlyLeuProGluuysGlyProlyProPhaArg 40
Db      138 AAATACGAAAGGAGCACCGAGCTGGGCTCCAGAGGCAAGGGGGCTTAAGCTTTTCGA 197

QY      41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db      198 AGCTACAAACAACAGCTGATCATTTGGGATTTGTATAGAGCGAGCTGCTCTCTG 257

QY      61 ThrAlaArgGluuAlaLysGlnIleArgArgGluuIleSerArgLysSerLysTyrValAsp 80
Db      258 ACTGCGGGGAGGCGAAGCAAAATTCCGGGAGATCAGCCGAAAGCAAGTGGGTGGAT 317

QY      81 MetLeuGluYAspTTPGluYArgTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db      318 ATCTGGGAGACCTGGGAAATACAAAGACAGCAAGAAAGCTCAATGATTCAGCGCTACAG 377

QY      101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
Db      378 GGAATGCCCATGAAACATCCGGGGCCGATGTGGTCACTCTCTGAAACATTGAGGAAATG 437

QY      121 LysLeuYAsnProGlyArgTyrGlnIleMetLysGluYArgLysSerSerGlu 140
Db      438 AAGTGAAGAAACCCCGGAAGATACAGATCATGAGAGGAAGGCAAGAGTCACTGAG 497

QY      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIleIlePheArg 160
Db      498 CACATCCAGCGCATCGACCCGGGACATMAAGCGGACATTAAAGAAAGCTATGTTCTTCAG 557

QY      161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db      558 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCGGCAATGAGAGTAT 617

QY      181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
Db      618 AACCGGAGGTGGGCTACTGCAAGGACCTGAGCCACATCCCGCTTGTCTCTCTAT 677

QY      201 LeuProGluGluAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db      678 CTTCCTGAGGAGGATGATCTGTGGGCACTGTGACAGCTGTGGCCAGTGAAGAGCACTCC 737

QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db      738 CTGCAAGGATTTCAAGCCCAAAATGGGGAGCCGTCCAGGGGCTCCAAAGCAAGAGAG 797

QY      241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db      798 CAGTGTGTAGCCACGTCAACAATCCAAAGACATATGGGGCATGAGACAAAGAAAGATCTATGT 857

QY      261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGly----- 276
Db      858 GGGCAATGTATCCCGTTAGGCTGCTCATCCGATATATTGATTGACGGGTTAAGAGAGCAT 917

QY      276 ----- 276
Db      918 AGGAGAACCTGGCTAGAGGACCTTCTTGCTCGAGTGCCTGTCTTCCCAAGCCGAG 977

QY      276 ----- 276

```

```

Db      978 GGTCTGCTCACTCCAGAGCCAGAGAGGCTCAGCGGGTCCCAAAAGACACACAAGCA 1037
QY      276 ----- 276
Db      1038 AAACCTCTGCCAAGAGGGGTATCCAGGGCAATGGCTGGGGCTCAGGCCAGCTCA 1097
QY      276 ----- 276
Db      1098 TGGGCACTGGGCCAGGACCCGACTTGAAGGGCTCAGGGAAAGCTCAAGCCCTGGGCA 1157
QY      276 ----- 276
Db      1158 AGCCCTCTCTCAGAGACCATCCCACTAAATAGTACCCCATGAGAGGCTTCA 1217
QY      276 ----- 276
Db      1218 AGACTTGTCTGACCCAGCGTCTGTGAAGGCTCAGCCGACTTATGGGGAAGTCACTG 1277
QY      276 ----- 276
Db      1278 ACTTGAAGACTAAGCCCAAGCTGCGAGCTCGAGCCACAGCCCACTGGAAAGCA 1337
QY      277 ----- 1SerLeuGlyLeuThrLeuArgLeuTyr 286
Db      1338 CCAAGTTCTTTCAACACTGCTGTCCCAAGATCTCTCGGGCTCACCTGGGCTGTG 1397
QY      286 PAspValTyrLeuValGluGlyGlnAlaLeuMetProIleThrArgIleAlaPheLys 306
Db      1398 GAGCGGTATCTGTGTGAAGCGCAACAGGGGTGATGCCATATACAAAGATCCCTTTAA 1457
QY      306 sValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTTPAlaArgPheCysAs 326
Db      1458 GGTTCACAGAACCGCTCACGAAGCTCAAGTGTGGCTGGCCCTGGGCACTTTTGCA 1517
QY      326 nArgPheValAspThrTTPAlaArgAspGluAspThrValLeuLysHisIleuArgAlaSe 346
Db      1518 CCGGTTCTGTGATACCTGGGCGAGGATGAGACATGTCTCAAGATCTTATGGGCTC 1577
QY      346 ThrLeuLysLeuThrArgLysGlnGlyLysAspLeuProProProAlaLysProGluGln 366
Db      1578 TATGAAGAAACTAAACAAGAGAGGGGGCTGACACCCCAAGCCCAAGCAGCAAG 1637
QY      366 YSerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAs 386
Db      1638 GTCTGGGACCTCAAGGCTGTGTGCGGCTTCACGTGGGGAAGACCTCTGCAAGGGGGA 1697
QY      386 PArgGlnAlaProProGlyProProAlaArgPheProArgProIleTTPSerAlaSerPr 406
Db      1698 CAGGAGGGCCCTCCAGGGCCCAACGCCGGTTCGGCGGCCCATTTGGTCACTTCCCT 1757
QY      406 oProAlaAlaProArgSerSerThrProCysProGlyGlyAlaValAlaArgGluAspThrTy 426
Db      1758 GCCACGGGCACTCGTTCTTCCACACCTGTCTGTGGGGCTGTCCGGGAAGACACTTA 1817
QY      426 rProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyProGlnGlySe 446
Db      1818 CCTGTGGGACTCAAGGGTGTGCCCCAGCCGGGCTCGGCTCAAGAGAACTTCAAGGGTTC 1877
QY      446 rTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThrAspLeuAspValGluGlu 466
Db      1878 CTGGAATTCCTGTCAATGGAATCCATGTGCCGCTCCCAAGGACTGGACCTTAAGGG 1937
QY      466 YProTTPPheArgHisTyrAspPheArgLysSerCysTyrValAlaArgAlaIleSerGlnGlu 486
Db      1938 CCTTGTGTCGCGCATTAATGATTGACAGAGAGCTGTGGGTCCGTGGCATATCCAGGA 1997
QY      486 uAspGlnLeuAlaProCysTTPGlnAlaGlnHisProAlaGluArgValArgSerLAsp 506
Db      1998 GAGCAGCTGGCCCCCTGTGGAGGTAAACCTCGGAGACGGGTTAGATCGGCTTT 2057
QY      506 eAlaAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGluInPr 526
Db      2058 CACTGCACCAGCACTGATTTCCGACCAAGGGCAACCCCTTCAGAGCTAAGGAGCAACAGCA 2117

```


QY 526 oCyAlaProthrsSerGlyProCybLeuCyAGlyLeuHisLeuGluSerSerGlnPhePr 546
DB 2118 GTGTGCTCCACCTCAGGCGCTTGTCTGTGGCGCTCCACTTGAGAAAGTTCTCAGTTCCC 2177
QY 546 oProGlyPhe 549
DB 2178 TCCAGGCTTC 2187
RESULT 11
AX775951 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775951
DEFINITION Sequence 221 from Patent WO03048202.
ACCESSION AX775951
VERSION AX775951.1 GI:32693669
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Matsuda, A. and Muramatsu, S.
AUTHORS NF-kB activating gene
TITLE Patent: WO 03048202-A 221 12-JUN-2003;
JOURNAL Aeshi Kasei Kabushiki Kaisha (JP)
FEATURES
Source Location/Qualifiers
1..2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11675.1"
/db_xref="GI:32693670"
/translation="MDVVEVAGSWMAQERED1IMKEKGRHAGLPEDKGPFRFSYNN
NVDHGI VHTHELPLTAREAKOIRREISRSKVMIDGMEKXSKLIDRAYKGM
PMNIRGIMMSVLTNTEEMKLNKRGROYIMKEKGRSSEHIOIRDVSGTLRKIIIPR
DRYGTOKREHLHTILAYEVNPEVGYCDLSHTALFLIYIPEDDAPALVQLASER
HSLOGFHSNGVITQGLQDQEHVATSRQPTWGHQISLGLTLMDVYLVEGSOLM
PIIRIAFKVQCKRLTKTSRCGPWAFPCNRFDVTWABEDVTLKILRSMKGLTKQGD
LPPEAKPEQGSASRPVPSRSGKTLCSGDRQAPGPAPRPRIIPSPAPRARSST
PCPGAVBEDDPVGTGVSRLAAGSGQSWRFLOMNSMPLPTDLDVSGPWRHY
DRRSQWRAISQBDOLAPCQAEHRAERVSAFAABETSDQETPRARDEQCART
SGPCLCGHLESQPPPGF"
ORIGIN
Alignment Scores:
Pred. No.: 1,586-89 Length: 2647
Score: 2835.00 Matches: 526
Percent Similarity: 95.88 Conservative: 0
Best Local Similarity: 95.88 Mismatches: 1
Query Match: 95.04 Indels: 22
DB: Gaps: 1
US-10-071-838-2 (1-549) x AX775951 (1-2647)
QY 1 MetAspValValGluValAlaGlySerTrpTAlaGlnGluArgGluAspIleLeuMet 20
DB 755 ATGAGCGTGGTAGAGGTGCGGGGCACTTGTTGGGCAACAAGCGAGACATCTATTATG 814
QY 21 LysTrpGlyLysGlyHisAlaGlyLeuProGluAspIleProGlyProLysProPheArg 40
DB 815 AAAATACGAAAGGACACCGAGCTGGCTGCCAGAGCAAGAGGGGCTTAAGCTTTTTCGA 874
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 875 AGCTACAAACAACAGCTGATCATTTGGGGATTGACATGAGACGAGACTGCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 935 ACTCGGGGAGGAGCGAAGCAATTGCGGGGAGATCAAGCCGAAAGACGAAGTGGGTGAT 994

QY 81 MetLeuGlyAspTrpGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 995 ATGCTGGAGACACTGGAGAAATACAAAGACAGAAAGCTCATAGTCCAGCTACAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 1055 GGAATGCCCATGAACATCCGGGGCCGATGTGTGACGTCTCTGGAACACTGAGGAATG 1114
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysArgSerSerGlu 140
DB 1115 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAAAGGCAAGAGTCACTCGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 1175 CACATCCAGCGCATCGACCGGGAGCTAAGCGGACATTAAGAGACATATTCTTCAGG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 1235 GATCGATACGGAACCAAGACAGCGGGAACCTACCATCTCTCGCATATGAGAGTAT 1294
QY 181 AsnProGluValGlyTyrCybArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
DB 1295 AACCCGAGGTGGCTACTGACAGGACCTTAGCCACATCCCGCTTGTTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 1355 CTTCCTAGAGAGATGATCTGTGGCACTGTGTGACGCTGTGGCAGATGAGGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnGlnGlu 240
DB 1415 CTGCAGGAGATTTCACACCCCAATGGCGGAGCCGTCCAGGGGCTCCAAAGACCAACAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCyb 260
DB 1475 CATGTGTAGCCAGCTACAAACCCAAACCAATGGGATCAG----- 1516
QY 261 GlyGlnCybSerProLeuGlyCybLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 1517 -----ATCTCTCTGGG 1558
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB 1529 CTGACCTCGGCGCTGTGGAGCTGTATCTGTAGAAAGCGCAAGCGTTGATCCGATA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysTrpSerArgCybGlyPro 320
DB 1589 ACAAGAAATCGCTTTAAGTTACAGCAGAAAGCGCTCACGAAGAGCTCAGGTGGCCG 1648
QY 321 TrpAlaArgPheCybAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1649 TGGGCACTTTTTCACACCGGTTGTTGATACCTCGGGCCAGGATGAGACACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 1709 AAGCATCTTAGGGCTCTATGAAAGAACTAAAGAAAGACAGGGGAGCTGCCACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerLysSerArgProValProAlaSerArgGlyLys 380
DB 1769 GCCAAACCCGAGCAAGAGGTGTCTCGCATCAAGGCTGTGCGGCTTACCGTGGGGAAG 1828
QY 381 ThrLeuCybLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1829 ACCCTCTGCAGAGGGGACAGGACAGGCCCTTCAGGCCACCAAGCCGGTCCCGGGGCC 1888
QY 401 IleTrpSerAlaSerProProAlaArgAlaProArgSerThrProCybProGlyGlyAla 420
DB 1889 ATTTGGTCAAGTTCCCGGCAAGCGGACCTGTTCTTCCACACCTGTCTCGTGGGGCT 1948
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1949 GTCCGGGAAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCGGGCTGTGCTCAG 2008

Qy 441 G1yG1yProGlnGlySerTrpArgPheLeuGlnTrpPheSerMetProArgLeuProThr 460
 Db 2009 GAGGAGCTCAGGGTCTCTGAGATTCTCGAGTGAAGTCACTCCAGCCCGCTCCCAAG 2068
 Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTrpAspPheArgGlnSerCysTrpVal 480
 Db 2069 GACCTGACCTAGAGGGCCCTCGGTCCGCATTATGATTACAGACAGAGCTGTGGATC 2128
 Qy 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTrpTrpGlnAlaGlnHisProAlaGln 500
 Db 2129 CGTGCCATATCCAGGAGGAGCACTGACCTGCCCCCTGCGCAGCTGAACACCCCTGCGAG 2188
 Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 Db 2189 CGGGTGAATCGGCTTTCGCTGCAACCACTGATTCGACACGGGCAACCCCTTACGA 2248
 Qy 521 AlaArgAspGlnGlnProCysValaProThrSerGlyProCysLeuGlyLeuHisLeu 540
 Db 2249 GCTAGGAGCAACAGCGGTGCTCCACCTCAGAGGCTTGCTGCTGCGGCTCTCACTTG 2308
 Qy 541 GluSerSerGlnPheProProGlyPhe 549
 Db 2309 GAAAGTCTCAGTTCCTCCAGGCTTC 2335
 RESULT 12
 AR578515
 LOCUS AR578515 691 bp DNA linear PAT 14-DEC-2004
 DEFINITION Sequence 691 from patent US 6783969.
 ACCESSION AR578515
 VERSION AR578515.1 GI:56581311
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2072)
 AUTHORS Tang, Y.T., Goodrich, R.W., Asundi, V. and Drmanac, R.T.
 TITLE Chapsin V-like polypeptides
 JOURNAL Patent: US 6783969-A 691 31-AUG-2004;
 Nuveio, Inc.; Sunnyvale, CA
 FEATURES
 source 1..2072
 /organism="Unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,36e-89 Length: 2072
 Score: 2827.00 Matches: 525
 Percent Similarity: 95.6% Conservative: 0
 Best Local Similarity: 95.6% Mismatches: 2
 Query Match: 94.8% Indels: 22
 Gaps: 1
 DB: 6
 US-10-071-838-2 (1-549) x AR578515 (1-2072)

Db 403 ATCTGGAGACTGGGAGAAATACAAAGCAGCAAGAACTCATAGTACGAGCTACAG 462
 Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuMetAsnIleGluGluMet 120
 Db 463 GGAATGCCATGAAACATCCGGGCGCATGTGTCACTCTCTGAACTGAGGAATG 522
 Qy 121 LysLeuLysAsnProGlyArgTrpGlnIleMetLysGluLysArgSerSerGlu 140
 Db 523 AAGTTAAAAACCCCGAAGATACCAATCATAGAGAGAGAGGAGCAAGAGTCACTGAG 582
 Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgHisIlePhePheArg 160
 Db 583 CACATCAGAGCATTCGACCGGAGCTGAAGGGGACATTAGAGAGCATATATCTTCAGG 642
 Qy 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGluTrp 180
 Db 643 GATCGATACGAAACCAAGCAGCGGAACTACTCACACTCTCTGCGATATGAGAGTAT 702
 Qy 181 AsnProGluValGlyTrpCysArgAspLeuSerHisIleAlaAlaPheLeuLeuTrp 200
 Db 703 AACCCGAGGTGGGCTACTGCAAGGACTGAGCAATCGCCGCTTCTCTCTTAT 762
 Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 763 CTTCGTAGAGAGATGATTCGTGGCACTGTGCAGCTGTGCGCAGTGAAGGCACTTC 822
 Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlu 240
 Db 823 CTCAGGAGATTTCAACGCCCAATATGGCGGACCGTCCAGGGGCTCCCAAGACCAACAGAG 882
 Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 883 CATGTGTACCACTGTCACCAACCAAGACCATGGGAGATG----- 924
 Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 925 -----ACTCTCTCTGGG 936
 Qy 281 LeuThrLeuArgLeuTrpAspValTrpLeuValGluGlyGluAlaLeuMetProIle 300
 Db 937 CTCACCTGCGCTGTGGGACGTGTATCTGTGTGAAGGCCCAACAGGGTGTGATGCCATA 966
 Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 997 ACAAGATCGCTTTAAAGTTACGACAGAGCGCTCAAGAGAGTCCAGTGTGGCCG 1056
 Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 1057 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGATAGAGCACTGTCTC 1116
 Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
 Db 1117 AACCATCTTAAAGGCTCTATGAAGAACTAAACAAAGAAAGCAGGGGCTGCAACCCCA 1176
 Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProIleSerArgGlyGlyLys 380
 Db 1177 GCCAAACCCGAGCAAGAGTGTCTGGCATTCAGGCTGTGTGCGGCTTCACTGTGCGGGAG 1236
 Qy 381 ThrLeuCysLeuGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1237 ACCCTTCGAAGGGGAGCAGGAGCCCTCCAGGCCCAACACCCGGTCCCGCGGCC 1296
 Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysAspProGlyVala 420
 Db 1297 ATTGTGTAGCTTCCCGGCCACCGGACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1356
 Qy 421 ValArgGluAspThrTrpProValGlyTrpGlnGlyValaProSerProAlaLeuAlaGln 440
 Db 1357 GTCCGGAGAGACCTTACCTGTGGGCACTCAGAGGTGTGCCAGCCGGGCTGTGGCTCAG 1416
 Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
 Db 1417 GAGAGACCTCAGGCTTCTGAGATTCCTGCAATGGAATCTCATGTCCCGCCGCCCAAG 1476

QY 461 AspleubapValaGluGlyProTrrPheArgHisTyrAapPheArgGlnSerCySeTrrVal 480
DB 1477 GACCTGACGATGAGAGGCCCTTGGTCCGCAATTATGATTTTCAGACAGAGCTGCGGATC 1536
QY 481 ArgaIaIesSerGlnGluAapGlnLeuAaProCySTrrGlnaGluHisProCaGlu 500
DB 1537 CGTCCATATCCACAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTCGGAG 1596
QY 501 ArgaIaIesSerAaIaPheAaIaProSerThrAapSerAapGlnGlyThrProPheArg 520
DB 1597 CGGATGAGATCGGCTTTCGTCGACCCGACCTGATTTCCAGCAGGAGCCCCCTTCAGA 1656
QY 521 AAlaArgAapGlnGlnProCyAaIaProThrSerGlyProCySeLeuCySeGlyLeuHisLeu 540
DB 1657 GCTAGGACGAAACACCCCTGCTCCACCTCAAGGCTTGCCTGCGGCTCCACATTC 1716
QY 541 GluSerSerGlnPheProGlyPhe 549
DB 1717 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1743
RESULT 13
AX775949 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS Sequence 219 from Patent WO03048202.
DEFINITION AX775949
ACCESSION AX775949
VERSION AX775949.1 GI:32693667
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Matsuda,A. and Muramatsu,S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 219 12-JUN-2003;
Aeshi Kasei Kabushiki Kaisha (JP)
FEATURES
source location/Qualifiers
1. 2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11674.1"
/translation="MDVVEVAGSWMAQEREDINKYBKGRAGLPEDKGPFRSYNN
NVDHGIIVHETELPPLTARBAKQIRREISRKSNVMDKMEKTSRKLTIDRAYKGM
PMNIRGPMWSVLTANTEBKLNKRGROYIMKEKGRSSHHIQRIDRVSGLRKHIIFR
DRYCTKORELHILILAYEVNPEVYSCDSHIALPLIYTPEDDAFWALVOLLASR
HSLDGFSPNGSIVQGLQDOQEHVAVTSQPTMGHQSILGTTLEMLNDVLYVEGQALM
PITRIAFKVOQKRLTKTSRCGPMARPCNRFVDTWARBDVTLKHLRASMKKLTQGD
LQPPAKPEOGSSASRPVPASRGKTLCGDRQAPPGPARPPIWASPERARSSST
PCPGAVEDTYPVGTGVSPPALAQGSPQSMFLQWNSMPLPTDLVGVPMFRHY
DFRSCWVRATISOEDOLAPCQAEHPARVARSAPASTSDQCTPRARBDGCAPT
SGPRLCGHLHSSQPPPGF"

ORIGIN

Alignment Scores:

Pred. No.: 2.99e-89 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.64 Conservative: 0
Best Local Similarity: 95.64 Mismatch: 2
Query Match: 94.84 Indels: 22
DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x AX775949 (1-2647)

QY 1 MetAapValaGluValaIaGlySerTrrPalaGlnGluAapIleIleMet 20
DB 755 ATGACGATGATGAGGTCGCGGCGAGTTGTGTGGCACAAGAGCAGAGCATCATTTATG 814

QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAapLysGlyProLysProPheArg 40
DB 815 AAATACGAAAGGAGACCCAGAGCTGGCTCCAGAGGACAAAGGGGCTTAAGCCTTTTCGA 874
QY 41 SerTrrAaAaAaAaValAaPheIleGlyTlleValHisGluThrGluLeuProLeu 60
DB 875 AGTACAAACAACAACGTCGATCTTTTGGGATTTGATGATAGCGAGCTGCCCTCTCG 934
QY 61 ThrAlaArgGluAaLysGlnIleArgArgGlnIleSerArgLysSerLysTrrValAap 80
DB 935 ACTGCGCGGAGGCGGAGCAAAATTCGCGGAGATTCAGCCGAAAGCCAAAGTGGTGGAT 994
QY 81 MetLeuGlyAapTrrGluLysTyrLysSerSerArgLysLeuIleAapArgAlaTyrLys 100
DB 995 ATGCTGGAGACCTGGAGAAATTCAAAGACGAGAAAGCTCATAGATTCAGGCTTACAG 1054
QY 101 GlyMetProMetAaIleArgGlyProMetTrrSerValLeuLeuAaIleGluGluMet 120
DB 1055 GGAATGCCATGAAACATCCGGGGCCGATGTGTCAGTCCCTCGAACACTGAGGAATG 1114
QY 121 LysLeuLysAaAaProGlyArgTyrGlnIleMetLysGluLysArgSerSerGlu 140
DB 1115 AAGTTGAAAAACCCGGAAGATACAGATCATGAGAGAAAGGCAAGAGTCACTGAG 1174
QY 141 HisIleGlnArgIleAapArgAapValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 1175 CACATCCAGGCGCATTCGACCGGAGCTGAGGGGACATTAAGAAACATATATCTTCAGG 1234
QY 161 AapArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 1235 GATGATGATCGGAAACCAAGCAGCGGAACTCTCACATCTCTCGGATATGAGAGTAT 1294
QY 181 AaAProGluValaGlyTyrCyAaAapLeuSerHisIleAaIaAaLeuPheLeuTyr 200
DB 1295 AACCGGAGGTTGGCTACTGAGGACCTGAGCACATCCCGCTTGTCTCTTAT 1354
QY 201 LeuProGluGluAaAaPheTrrPalaLeuValGlnLeuLeuAaSerGluArgHisSer 220
DB 1355 CTTCCTGAGGAGATGATCTGCGGCACTGTGAGCTGCTGGCCATGAGAGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAaGlyGlyThrValGlnGlyLeuGlnAaGlnGln 240
DB 1415 CTGCAAGGATTTTCACACCCAAATGGCGGAGCCCTCCAGGGGCTCCCAAGCAACAGAG 1474
QY 241 HisValValaIaAaThrSerGlnProLysThrMetGlyHisGlnAaLysLysAaPheCys 260
DB 1475 CATGTGTAGCCACGTACAACCCAAAGCATGGGGCATCAG----- 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAaPglyIleSerLeuGly 280
DB 1517 -----ATCTGCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrrAapValTyrLeuValGluGlyGlnIaAaLeuMetProIle 300
DB 1529 CTCACTCTGGCGCTGTGGAGCTGTATCTGTAGAAAGCCAAACAGGCGTGAATGCCATA 1588
QY 301 ThrArgIleAaIaPheLysValaGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 1589 ACAAGAAATCCCTTAAAGTTACAGCAAGAGGCTTCAGAAAGATCCAGATGTGGCCG 1648
QY 321 TrrPalaArgPheCysAaAaPheValAaPThrTrrPalaArgAapGluAapThrValLeu 340
DB 1649 TGGGCAAGTTTTCGAAACCGGTCGTGTAACCTGGGCGAGGATGAGACATGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAaLeuProProPro 360
DB 1709 AAGCATCTTAGGGCTCTATGAAGAACTAAAGAAAGACAGAGGCGGACTCTGCAACCCCA 1768
QY 361 AAlaAapProGluGlnGlySerSerAaAaSerArgProValProAlaSerArgGlyLys 380
DB 1769 GCCAAACCCAGACAGAGGATGTGCGGCATTCAGAGGCTGTGCGCTTCACGTGGGGAGG 1828

QY 381 Thleucysylgyslaasparaglnalaproproglyproproalaaargpheaargpro 400
 DB 1829 ACCCTTCGAAGGGGAGACAGGCAAGCCCTTCAGAGCCCAAGCCGGTTCGCCGCCCC 1888
 QY 401 lletfPserlaserproproalalaproargserthrprocyasproglyvala 420
 DB 1889 ATTGGTCACCTTCCCGCCACGCGGCACTCGTTCCTTCACACCTGCTCGTGGGGCT 1948
 QY 421 Valarglunapthrrtyrprovalglthyglnlvalproserproalaleuallgln 440
 DB 1949 GTCCGGAGAGACACTTACCTGTGGGACCTAGGGGTGTGCCAGCCGCGCTGCTAG 2008
 QY 441 GlyglYproglnglyserthrparpheleuglntirpansermecproargleuprothr 460
 DB 2009 GAGAGACCTCAGGGTCTCTGGAGATTCTGTGAGATGGAATCTCAAGCCCGCTCCCAAG 2068
 QY 461 Aspleuapvalgluglyprotrpheaarghietryrappheaarglinserytrpval 480
 DB 2069 GACCTGGACGTAGAGGGCCCTTGCTCCGCAATATGATTTCAAGACAGAGCTGTGGGTC 2128
 QY 481 ArgAlalIeserlnglunaprglnleuallaprocyertrpnlalaglunleproalglu 500
 DB 2129 CGTGCCATATCCAGAGAGACAGCTGCCCCCTGTGGCAGGCTGAACACCTGCGAG 2188
 QY 501 ArgValaArgserAlaPheAlaAlaPserThrApsersaprglnglythrproPheaArg 520
 DB 2189 CGGGTGATTCGGCTTCGCTGCACCCAGCACTGATTCGACCGGGCAACCCCTTCAGA 2248
 QY 521 AlaaArgapgluglunprocyasalaprothrserglyprocyaleuHleuHleu 540
 DB 2249 GCTAGGAGACAAAGCCGTGTGCTCCACCTCAGGGGCTTGCTGCGGCTCACTTG 2308
 QY 541 GluserSerlnPhePproglYpHe 549
 DB 2309 GAAAGTCTCAGTTCCTCCAGGCTTC 2335

RESULT 14

AX833552 2647 bp DNA linear PAT 15-DEC-2003
 LOCUS Sequence 676 from Patent EP1347046.
 DEFINITION AX833552
 ACCESSION AX833552 GI:39919687
 VERSION AX833552.1
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE

1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
 Yamamoto,J.T., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
 Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
 Masuno,Y.
 Full-length cDNA sequences

TITLE

Patent: EP 1347046-A 676 24-SEP-2003;

JOURNAL

Research Association for Biotechnology (JP)

FEATURES

source location/Qualifiers
 1..2647 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 2,99e-89 Length: 2647
 Score: 2827.00 Matches: 525
 Percent Similarity: 95.6% Conservative: 0
 Best Local Similarity: 95.6% Mismatches: 2
 Query Match: 94.8% Indels: 22
 DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x AX833552 (1-2647)

QY 1 MetAspValalglunallaglYsertrpAlaglglunarglunaprglIleMet 20
 DB 755 ATGACCTGTTAGAGGTCCGGGACGTTGGTGGACAAAGAGAGACATCATTTATG 814
 QY 21 lyeTYglulYsglYhlsargalaglYleuproglunaprglYprolyPropheArg 40
 DB 815 AAATACGAAAGGAGACCGAGCTGGGCTGCAGAGAGCAAGGGGCTTAAGCCTTTTGA 874
 QY 41 SerTYrApsnApsnValApsnHlsleuglylIeValHlsgluthrgluleuproPheleu 60
 DB 875 AGCTACAAACAACAGCTGATTCATTTGGGATTTGATCATAGACGAGGCTGCTCTG 934
 QY 61 ThrAlaArgglunallaysglnlIeArgargglunlIeserApsYserlYstrpValAsp 80
 DB 935 ACTGCGGGAGGCGAAGCAAAATTCGGGGGAATCAGCCGAAAGCAAGTGGTGGAT 994
 QY 81 MetleuglYAspTrpglulYsTYrlyserSerArglylsleuIleAsparAlaTYrlyls 100
 DB 995 ATCTGGGAGACTGGGAGAAATCAAAAGCAGCAAAAGCTCATATGAGGCTACAG 1054
 QY 101 GlyMetPromeLanlIeArgglyPromeTrpSerValleuLeuAnlIeglulMet 120
 DB 1055 GGAATGCCATGAACATCCGGGGCCGAGTGTGTCAGTCTCTCGAAACCTGAGAAATG 1114
 QY 121 lYleuLylYAsnProglYargTYrYrnlIeMetYglulYsglylsAArgSerSerGlu 140
 DB 1115 AAGTTGAAAAACCCGGAAGATACAGATCATAGAGAGAGGCAAGAGTATCTGAG 1174
 QY 141 HlsIeglunarglIeAspArgAspValIserglyThrleuarghlsIlePhePheArg 160
 DB 1175 CATATCAGAGGACATGACCGGGACGTAAAGGGGACATTAAAGAGCATATATCTTCAGG 1234
 QY 161 AsparTYrTYrThrlysglnarggluleuLeuHlsIleleuLeuAlaTYrglulYr 180
 DB 1235 GATCGATACGGAACCAAGCAGCGGGAATCTCAATCCTCTCGCATATGAGAGTAT 1294
 QY 181 AsnProglunValglYrtyrCysArgAspLeuSerHlsIleAlaIlePheleuLeuTYr 200
 DB 1295 AACCCGAGGTGGCTACTGACAGGAGCTGAGCAATGCGCCCTGTGCTCTCTCAT 1354
 QY 201 leuproglulunaprglApsnAlaPheTrpAlaleuValglleuLeuAlaSerglunArgHlsSer 220
 DB 1355 CTTCCTGAGAGAGATCATTCCTGGGCACTGTGCACTCTGGCCAGTGAAGGACTCTC 1414
 QY 221 leuGlnglyPheHlsSerProanglnglylYthrValGlnglyleuGlnApsnglngln 240
 DB 1415 CTGACAGGATTTCAAGCCCAATGGCGGACGCTCAGAGGGCTCCAAAGCCAAAGGAG 1474
 QY 241 HlsValValAlaThrserGlnProlysthMetGlyHlsGlnApslylsAspLeuCys 260
 DB 1475 CATGTGTGTGCCACGTCAACCAAGCAATGGGCACTAG----- 1516
 QY 261 GlyGlnCysSerProleuglyCysleuIleArgIleleuIleApslylIeserleugly 280
 DB 1517 -----ATCTCTCTCGGG 1528
 QY 281 leuThrleuArgleuTrpAspValTYrleuValgluglyglunAlaleuMetProle 300
 DB 1529 CTCACCTCGCGCTGTGGAGCGTGTATCTGTAGAAAGGGAACAGCGTTGATGCCGATA 1588
 QY 301 ThrArgllleAlaPheIysValGlnglnlysaArgleuThrlYethSerArgYsglyPro 320
 DB 1589 ACAAGATGCGCTTTAAGGTTAGCAGAAAGCCCTCACAAAGACGTCCAGGTGTGCCCG 1648
 QY 321 TrpAlaArgPheCysAsnArgPheValAspTrpTrpAlaArgAspglunaprglYlleu 340
 DB 1649 TGGGCACTTTTTCACACCGGTTCTGTATACCTGGCCAGAGATGAGCACTGTGCTC 1708
 QY 341 lYsHlsleuArgAlaSerMetlylsYsleuThrArglysglnglyAspLeuProProPro 360
 DB 1709 AAGCATCTTAAGGCGCTCTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCACACCCCA 1768
 QY 361 AlalyPproglunlnglyserSerlaserArgprovalProAlaSerArgglylYlys 380

```

Db      1769 GCCAAACCCGAGCAGAGGCTGCTCGCATCCAGGCGCTGTGCGGGCTTCACTGGCGGGAG 1828
Qy      381 Thrleucylslysglyaaparglnalalaproproglyproproalaaargpharproargpro
Db      1829 ACCCTTCGCAAGGGGAGCAGGAGCCCTCCAGGCCACAGCCGGGTTCCCGGGCCC 1888
Qy      401 lletpseralaserproproargalalaproargserthprocyargglyalal 420
Db      1889 ATTGGTCAGCTTCCCCCGCACGGGACCTGCTTCCACACCCTGTGCTGGGGGCT 1948
Qy      421 Valarglualapthrrtyrprovalglythrsinglyvalproserproalaleualagln 440
Db      1949 GTCCGGGAGACACTACCTGCTGGGCACTCAGGGGTGGCCAGCGCCGCGCTGGCTAG 2008
Qy      441 Glylylproglnglysertrparaphleuglntfpansermetproargleuprothr 460
Db      2009 GGAGGACCTCAGGGGTCTCTGAGATTCTGCAAGTGGAACTCCATGCCCGCTCCCAAG 2068
Qy      461 Aspleuapvalaglgllyprotrpthearghistryaaptheargglinsercyatrval 480
Db      2069 GACCTGAGCTAGAGGGCCCTTGGTCCGCCATTATGATTTTCAGACAGAGCTGCTGGGTC 2128
Qy      481 Arghalalieserlnglualapglaleualaprocytrpgrlnalaglulhlaproalaglu 500
Db      2129 CGTGCATATCCAGAGAGACAGCTGCGCCCTGCTGGAGGCTGAACACCTGCGGAG 2188
Qy      501 Argvalarseralalaphenalalaproserthrapserrarpglnglythpropharg 520
Db      2189 CGGGTGTGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACAGAGGAGCCCTTCGA 2248
Qy      521 Alaaraparglulglnprocyalalaprothserglyprocyaleucyaglyleuhsilenu 540
Db      2249 GCTTGGAGCAACAGCCCTGTGCTCCCACTCAAGGCTTGTGCTGCGGCTCCACTTG 2308
Qy      541 GluBerserglnpheproproglyphe 549
Db      2309 GAAAGTCTCAGTTCCTCCAGGCTTC 2335

RESULT 15
AK095385 2647 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ38066 f18, clone CTONG2015330, moderately
DEFINITION similar to H. sapiens mRNA for tre oncogene.
ACCESSION AK095385.1 GI:21754631
VERSION AK095385.1
KEYWORDS oligo capping, f18 (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Iehii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsube, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimura, M., Watanabe, M.,
Hiraoa, S., Chiba, Y., Iehida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hottuta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,
Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuki, F., Hara, R.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, A.,
Shichita, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terasima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,

```

```

Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Oimori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togsaki, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, N., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuo, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
2
JOURNAL PubMed
REFERENCE 14702039
AUTHORS Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M.,
Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Iehii, S.,
Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuo, Y., Nagai, K. and Isogai, T.
NEBO human cDNA sequencing project
TITLE NEBO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2647)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatacari, Katsuzu, Chiba 292-0812, Japan
(B-mail: genomc@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
SOURCE location/Qualifiers
1..2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2015330"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Cloning vector: pMR18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 2,998-89 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 95.6% Mismatches: 2
Query Match: 94.8% Indels: 22
DB: Gaps: 1
US-10-071-838-2 (1-549) x AK095385 (1-2647)
Qy      1 Metapvalalglualalaglalysertrprralaglnlualarglualapllellec 20
Db      755 ATGACGCTGTGATGAGCTCGCGGCGAGTGTGGGCAACAAGAGCAGAGCATCATATATG 814
Qy      21 Lyeftrglulysglyhlsargalagllyleuproglunbrlysglyprolyarpropharg 40
Db      815 AAATACCAAAAGGACACCGAGCTGGGCTGCAGAGCAAGAGGCGCTTAAGCTTTTCGA 874
Qy      41 SerTyranbapnaenValaaphlaLeuglyllevAlhlsigluthrgluleuproloeu 60
Db      875 AGCTACCAACAACGCTGCATCATTTGGGATTTCTATAGACGAGAGTCCCTCCTCTG 934
Qy      61 Thrllaarglualalalyglnlileargarglulleserarglyserlystrpvalaap 80

```

Db 935 ACTGCGGGAGGCGAACAATTGGCGGAGATCAGCCGAAGCAAGTGGGTGAT 994
 Qy 81 MetLeuGlyAspTrpGlySerTyrLeuSerSerTyrLeuLeuLeuAlaTyrLeu 100
 Db 995 ATGCTGGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCATGATTCAGCTTACAG 1054
 Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
 Db 1055 GGAATGCCCATGAAACATCCGGGGCCCATGTGTGTCATCTCTGAAACATGAGAAATG 1114
 Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 Db 1115 AAGTTGAAAAACCCCGGAGATACCATGATCATGAAAGAGAGCAAGAGGTCTCTGAG 1174
 Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 1175 CACATTCAGCGCATTCGACCGGAGATGAGCGGACATTAAGAGACATAATCTTCAGG 1234
 Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
 Db 1235 GATCGATACCGAACCAGACGCGGAACTACTCCACATCTCTGCGATTTAGAGATAT 1294
 Qy 181 AsnProGluValGlyTyrCyAspArgLeuSerHisIleAlaIleLeuPheLeuTyr 200
 Db 1295 AACCCGAGGTGGGTACTGACAGGACCTGAGCCATCCGCCCTTCTCTCTAT 1354
 Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 1355 CTTCCTGAGAGAGATGATCTGCGGCACTGTGACGTGCTGCGCATGTAGAGGCACTCC 1414
 Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlnLeuGlnAspGlnGlnGlu 240
 Db 1415 CTGCGAGGATTTTCAACGCCCAATGCGGAGCGTCCAGGGGCTCCAAAGCAACAGGAG 1474
 Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 1475 CATGTGGTACCAAGTCAACAACCAAGCATGGGTCATCAG----- 1516
 Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 1517 -----ATCTCTCTCGGG 1528
 Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlnAlaLeuMetProIle 300
 Db 1529 CTACACCTCGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAACAAGCGTGTATGCCATA 1588
 Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 1589 ACAAGATGCGCTTTAAGTTACGACGAGAACGCGCTCAAGAGACGTCCAGGTGTGGCCG 1648
 Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 1649 TGGGCAAGTTTGGCAACCGGTTGTTGATTAACCTGGGCGAGGATGAGACACTGTCTC 1708
 Qy 341 LysHisLeuArgLysSerMetLysLysLeuThrArgLysGlnGlnLysAspLeuProPro 360
 Db 1709 AAGCATCTTAAGGCTCTATGAAGAACTAAACAAGAAAGCAAGGAGCACTGCAACCCCA 1768
 Qy 361 AlaLysProGluGlnGlnLysSerSerAlaSerArgProValProAlaSerArgGlyLys 380
 Db 1769 GCCAAACCGAGCAAGGTGTGTGCGCATCCAGGCTGTGCGGCTTCACTGGCGGGAG 1828
 Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1829 ACCCTCGCAAGGGGAGACGAGGCGCTCCAGGCCCAACAGCGCGTTCCCGCGGCC 1888
 Qy 401 IleTrpSerAlaSerProArgAlaProArgSerSerThrProCysProGlyValAla 420
 Db 1889 ATTGTGACGTTTCCCGGCAAGGCACTGTGTTCTTCAACCTGTCTGTGGGGCT 1948
 Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1949 GTCCGGGAAGACCTACCTGTGTGGGACCTCAGGGTGTGCTCAGCCCGGCTGGCTCAG 2008

Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
 Db 2009 GAGAGACCTCAGGGTCTCTGGAGATTCCTGCAATGGAATCCATGCCCCCTCCCAAG 2068
 Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgLysSerCysTrpVal 480
 Db 2069 GACCTGACCTAAGGGCCCTTGTGTTCGTCATTAATGATTGACAGACAGACTGTGGTTC 2128
 Qy 481 ArgAlaIleSerGlnGlnLeuAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
 Db 2129 CGTGCATATCCAGAGAGACAGCTGCCCCCTGTGCGAGCTGAAACACCTGCGGAG 2188
 Qy 501 ArgValArgSerLysPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 Db 2189 CGGTGAGATGGGCTTCTGCTGCAACCAAGCATGATTCGACCAAGGAGCAACCCCTTCAGA 2248
 Qy 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
 Db 2249 GCTAGGAGCAACAGCGGTGTGCTCCACCTCAGGGCTTGTCTGTGGGCTTCACTTG 2308
 Qy 541 GluSerSerGlnPheProProGlyPhe 549
 Db 2309 GAAAGTCTCAGTTCCCTCCAGGCTTC 2335

Search completed: April 5, 2006, 16:34:43
 Job time : 6744 sec

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 14:01:23 ; Search time 1619 Seconds
(without alignments)

10031.537 Million cell updates/sec

Title: US-10-071-838-1

Perfect score: 1964

Sequence: 1 atgacgctgctgagctgcgc.....aaatgtctctctgtgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1964	100.0	1964	US-10-071-838-1	Sequence 1, Appli
2	1929.8	98.3	1993	US-10-839-882-31	Sequence 31, Appli
3	1836.4	93.5	2084	US-10-302-172-692	Sequence 692, App
4	1812.4	92.3	2072	US-10-302-172-691	Sequence 691, App
5	1807.4	92.0	2647	US-10-108-260A-676	Sequence 676, App
6	1794	91.3	2146	US-10-071-838-3	Sequence 3, Appli
7	1750	89.1	1862	US-10-071-838-5	Sequence 5, Appli
8	1638.8	83.4	1752	US-10-094-466-61	Sequence 61, Appli
9	1220	62.1	7856	US-10-450-763-21335	Sequence 21335, A
10	1220	62.1	7856	US-10-450-763-24017	Sequence 24017, A
11	1220	62.1	7878	US-09-962-436-582	Sequence 582, App
12	1220	62.1	7878	US-10-843-641-1021	Sequence 1021, Ap
13	1220	62.1	7878	US-10-956-157-1259	Sequence 1259, Ap
14	999	50.9	8180	US-10-450-763-24016	Sequence 24016, A
15	999	50.9	8201	US-10-283-975A-272	Sequence 272, App
16	999	50.9	8201	US-10-278-698-640	Sequence 640, App
17	999	50.9	8201	US-10-278-698-1037	Sequence 1037, Ap
18	999	50.9	8284	US-10-450-763-21339	Sequence 21339, A
19	999	50.9	8408	US-10-098-841-51	Sequence 51, Appli
20	999	50.9	8420	US-10-098-841-50	Sequence 50, Appli
21	882.8	44.9	9804	US-09-764-869-1829	Sequence 1829, Ap
22	882.8	44.9	9804	US-09-764-891-7861	Sequence 7861, Ap
23	882.8	44.9	9804	US-10-091-504-1829	Sequence 1829, Ap

24	882.8	44.9	9804	US-10-227-577-1829	Sequence 1829, Ap
25	879.6	44.8	9404	US-09-764-869-1826	Sequence 1826, Ap
26	879.6	44.8	9404	US-09-764-891-7859	Sequence 7859, Ap
27	879.6	44.8	9404	US-10-091-504-1826	Sequence 1826, Ap
28	879.6	44.8	9404	US-10-227-577-1826	Sequence 1826, Ap
29	878	44.7	886	US-09-764-869-1832	Sequence 1832, Ap
30	878	44.7	886	US-10-091-504-1832	Sequence 1832, Ap
31	878	44.7	886	US-10-227-577-1832	Sequence 1832, Ap
32	876.4	44.6	886	US-09-764-869-1831	Sequence 1831, Ap
33	876.4	44.6	886	US-10-091-504-1831	Sequence 1831, Ap
34	876.4	44.6	886	US-10-227-577-1831	Sequence 1831, Ap
35	876.4	44.6	6696	US-09-764-869-1821	Sequence 1821, Ap
36	876.4	44.6	6696	US-09-764-891-7856	Sequence 7856, Ap
37	876.4	44.6	6696	US-10-091-504-1821	Sequence 1821, Ap
38	876.4	44.6	6696	US-10-227-577-1821	Sequence 1821, Ap
39	876.4	44.6	9805	US-09-764-869-1824	Sequence 1824, Ap
40	876.4	44.6	9805	US-09-764-891-7857	Sequence 7857, Ap
41	876.4	44.6	9805	US-10-091-504-1824	Sequence 1824, Ap
42	876.4	44.6	9805	US-10-227-577-1824	Sequence 1824, Ap
43	864.4	44.0	9796	US-09-764-869-1825	Sequence 1825, Ap
44	864.4	44.0	9796	US-09-764-891-7858	Sequence 7858, Ap
45	864.4	44.0	9796	US-10-091-504-1825	Sequence 1825, Ap

ALIGNMENTS

RESULT 1
US-10-071-838-1
; Sequence 1, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-00-610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1650)
US-10-071-838-1
Query Match 100.0%; Score 1964; DB 5; Length 1964;
Best local similarity 100.0%; Pred. No. 0;
Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGCTGTAAGCTGCGGAGATTGTGGGCAACAAGCCAGAGACATCAATTATG 60
DB 1 ATGACGCTGTAAGCTGCGGAGATTGTGGGCAACAAGCCAGAGACATCAATTATG 60
QY 61 AAATACGAAAGGACACCGAGCTGCGTCCAGAGGACCAAGGGGCTTAACCTTTTGA 120
DB 61 AAATACGAAAGGACACCGAGCTGCGTCCAGAGGACCAAGGGGCTTAACCTTTTGA 120
QY 121 AGTACCAACAACAAGCTGATCATTTGGGATTGTGATGAGACGAGCTGCTCTGTG 180
DB 121 AGTACCAACAACAAGCTGATCATTTGGGATTGTGATGAGACGAGCTGCTCTGTG 180
QY 181 ACTGGCGGAGGAGGAAAGCAATTTCGGCGGAGATCAAGCCGAAGAAGACAGTGCTGAT 240

Db	181	ACTGCGCGGAGGCGGAGCAAAATTGGCGGGAGATCAGCCGAAAGCAAGTGGGTGAAT	240
Qy	241	ATGCTGGGAGACTGGAGAAATACAAAGCAGCAAAAGCTCATGATCGACGGTAAAG	300
Db	241	ATGCTGGGAGACTGGAGAAATACAAAGCAGCAAAAGCTCATGATCGACGGTAAAG	300
Qy	301	GGAAATGCCATGATACATCTCGGGGGCCGATGTGGTCACTCTCTTGACATTTGAGGAATG	360
Db	301	GGAAATGCCATGATACATCTCGGGGGCCGATGTGGTCACTCTCTTGACATTTGAGGAATG	360
Qy	361	AAGTTGAAAAAACC CGGAGAGATACAGATCATGAAGGAGAGAGGCAAGGCTCATGTAG	420
Db	361	AAGTTGAAAAAACC CGGAGAGATACAGATCATGAAGGAGAGGCAAGGCTCATGTAG	420
Qy	421	CAATTCAGCGCATCGACCGGGACGTAAAGCGGGACATTTAAGGAACATATATTCTTCAGG	480
Db	421	CAATTCAGCGCATCGACCGGGACGTAAAGCGGGACATTTAAGGAACATATATTCTTCAGG	480
Qy	481	GATTCGATCGGAACCAAGACGCGGGAATACTTCACATCTCTCTGGCATATGAGGATAT	540
Db	481	GATTCGATCGGAACCAAGACGCGGGAATACTTCACATCTCTCTGGCATATGAGGATAT	540
Qy	541	AACCCGGAGTGGGCTACTGCAAGGGACCTGAAGCCCATATCGCCGCTTGTCCTCTCAT	600
Db	541	AACCCGGAGTGGGCTACTGCAAGGGACCTGAAGCCCATATCGCCGCTTGTCCTCTCAT	600
Qy	601	CTTCTCGAGAGAGATGCAATTCGGGCACTGGTGCAAGCTGCTGGCAGTGAAGGCACTCC	660
Db	601	CTTCTCGAGAGAGATGCAATTCGGGCACTGGTGCAAGCTGCTGGCAGTGAAGGCACTCC	660
Qy	661	CTGCAAGGATTTCAAGCCCAATGGCGGGACCGTTCAGAGGCTCTCAAGACCAACAGAG	720
Db	661	CTGCAAGGATTTCAAGCCCAATGGCGGGACCGTTCAGAGGCTCTCAAGACCAACAGAG	720
Qy	721	CATGTGTAGCCACGTGCACAAACCAGAACATGGGGCATCAGGACAAAGAAAGTCTAAGT	780
Db	721	CATGTGTAGCCACGTGCACAAACCAGAACATGGGGCATCAGGACAAAGAAAGTCTAAGT	780
Qy	781	GGGCAAGTGTTCGCCGTAGAGGCTGCTCATCCGGAATATTGAAGGAGCTCTCTCGGG	840
Db	781	GGGCAAGTGTTCGCCGTAGAGGCTGCTCATCCGGAATATTGAAGGAGCTCTCTCGGG	840
Qy	841	CTCACCCCTGCGCCTGTGGGACGTGTATCTGTAGAAGGCGAACAGGCGTTGATGCGATA	900
Db	841	CTCACCCCTGCGCCTGTGGGACGTGTATCTGTAGAAGGCGAACAGGCGTTGATGCGATA	900
Qy	901	ACAAAGATGCGCTTTAAGGTTCAAGCAAGAGCCCTTCAGAGAGCTCCAGGTGTGGCCCG	960
Db	901	ACAAAGATGCGCTTTAAGGTTCAAGCAAGAGCCCTTCAGAGAGCTCCAGGTGTGGCCCG	960
Qy	961	TGGGACGTTTTTGCAACCGGATTCGTGATACCTGGGCCAGGGATGAGGACATCTGTCTC	1020
Db	961	TGGGACGTTTTTGCAACCGGATTCGTGATACCTGGGCCAGGGATGAGGACATCTGTCTC	1020
Qy	1021	AAGCATCTTAAGGACCTCTATGAAGAACTTAACAAGAACAGGGGGACCTGTGACCCCA	1080
Db	1021	AAGCATCTTAAGGACCTCTATGAAGAACTTAACAAGAACAGGGGGACCTGTGACCCCA	1080
Qy	1081	GCCAAACCCGAGCAAGAGGCTGTGCGCATCAAGGCTGTGCGGCTTCACTGTGCGGAAAG	1140
Db	1081	GCCAAACCCGAGCAAGAGGCTGTGCGCATCAAGGCTGTGCGGCTTCACTGTGCGGAAAG	1140
Qy	1141	ACCCTCTGCAAGGGGGACAGGAGGCCCTCCAGGCCCAACAGCCCGGTTCCCGGAGGCC	1200
Db	1141	ACCCTCTGCAAGGGGGACAGGAGGCCCTCCAGGCCCAACAGCCCGGTTCCCGGAGGCC	1200
Qy	1201	ATTGTGTAGCTTCCCGGCAAGGAGCACTCGTTCTTCCACACCTGTCTGTGTGGGCT	1260
Db	1201	ATTGTGTAGCTTCCCGGCAAGGAGCACTCGTTCTTCCACACCTGTCTGTGTGGGCT	1260
Qy	1261	GTCGCGGAAAGACCTTACCTGTGGCACTCAAGGTTGTGCCAGCCGAGCCCTGTGCTAG	1320

Db 1261 GTCGGGGAAGAACACTACCTCGTGTGGGCACTCAGAGGTGTGGCCAGGCCGCTGTGGCTAG 1320

Qy 1321 GGAGGACCTCAGGGGTTCTCTGGAGATTCTCGCAGTGGAACTCCATGTCCCGCTCCCAAG 1380

Db 1321 GGAGGACCTCAGGGGTTCTCTGGAGATTCTCGCAGTGGAACTCCATGTCCCGCTCCCAAG 1380

Qy 1381 GACCTGGACGTAGAGGGGCTTTGTGTTCGGCAATTATGATTTTCAGACAGAGCTGTGGGTC 1440

Db 1381 GACCTGGACGTAGAGGGGCTTTGTGTTCGGCAATTATGATTTTCAGACAGAGCTGTGGGTC 1440

Qy 1441 CGTGCATATCCCAGAGAGACAGCTGGCCCCCTCTGTGCAAGGCTGAAACACCTCTGGAG 1500

Db 1441 CGTGCATATCCCAGAGAGACAGCTGGCCCCCTCTGTGCAAGGCTGAAACACCTCTGGAG 1500

Qy 1501 CGGGTAGATCGGCTTTGCTGTCACCCAGCATGATTTCCGACAGGGCAACCCCTTCAGA 1560

Db 1501 CGGGTAGATCGGCTTTGCTGTCACCCAGCATGATTTCCGACAGGGCAACCCCTTCAGA 1560

Qy 1561 GCTAGGAGCGAAGACCCGTGTGTCTCCACCTCAAGGCTTGTGCTGTGGGCTCCACTTG 1620

Db 1561 GCTAGGAGCGAAGACCCGTGTGTCTCCACCTCAAGGCTTGTGCTGTGGGCTCCACTTG 1620

Qy 1621 GAAAGTTCTCAGTTCCCTCCAGAGCTTCTAGAGCATCTGGGCGAGGCTCATGGCTGAT 1680

Db 1621 GAAAGTTCTCAGTTCCCTCCAGAGCTTCTAGAGCATCTGGGCGAGGCTCATGGCTGAT 1680

Qy 1681 AATTTCCCTAGGCTTAAACACCAGCAAGCTTGGCGTCTCTGTTTATTTTGTGTAA 1740

Db 1681 AATTTCCCTAGGCTTAAACACCAGCAAGCTTGGCGTCTCTGTTTATTTTGTGTAA 1740

Qy 1741 CTTATGAAAATGTATTAGAAAAGATGCACTCCGAGAGATTCAGAGATGGAACACACC 1800

Db 1741 CTTATGAAAATGTATTAGAAAAGATGCACTCCGAGAGATTCAGAGATGGAACACACC 1800

Qy 1801 AGACCCCGAGATCAAAAGCCAAACCATGCCCCGCTCCAGACACCCCGAGCCCAAGAC 1860

Db 1801 AGACCCCGAGATCAAAAGCCAAACCATGCCCCGCTCCAGACACCCCGAGCCCAAGAC 1860

Qy 1861 CATCGTCTGATTTCTGACGACACCGTGAGCTGTGCTTTGATCTTTAACTCATGAGAG 1920

Db 1861 CATCGTCTGATTTCTGACGACACCGTGAGCTGTGCTTTGATCTTTAACTCATGAGAG 1920

Qy 1921 ATTAACCTTCAAGCTTTTGAATAATGTTTCTCTGTGAATG 1964

Db 1921 ATTAACCTTCAAGCTTTTGAATAATGTTTCTCTGTGAATG 1964

RESULT 2

US-10-839-882-31

Sequence 31, Application US/10839882

Publication No. US20040203106A1

GENERAL INFORMATION:

APPLICANT: INCYTE PHARMACEUTICALS, INC.

APPLICANT: TANG, Y. Tom

APPLICANT: YUS, Henry

APPLICANT: HILIMAN, Jennifer L.

APPLICANT: GUEGLER, Karl J.

APPLICANT: CORLEY, Neil C.

APPLICANT: LAL, Preeti

APPLICANT: AZIMZAI, Yalda

APPLICANT: BAUGHN, Mariah R.

APPLICANT: JUNMING, Yang

APPLICANT: SHIH, Leo L.

TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS

FILE REFERENCE: PF-0619 PCT

CURRENT APPLICATION NUMBER: US/10/839, 882

CURRENT FILING DATE: 2004-05-05

PRIOR APPLICATION NUMBER: US/09/807,452

PRIOR FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 09/1175,737; unassigned; 60/118,559; 09/249,740; unassigned; 60/154,336

PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11; 1999-04-22

NUMBER OF SEQ ID NOS: 44
SOFTWARE: PERL Program
SEQ ID NO 31
LENGTH: 1993
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 4325626CB1
US-10-839-882-31

Query Match 98.3%; Score 1929.8; DB 8; Length 1993;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 ATGGAGCTGTAGAGTCCGGGAGTTGGTGGGCAACAGAGCGAGAGCATCATTTATG 60
DB 42 ATGGAGCTGTAGAGTCCGGGAGTTGGTGGGCAACAGAGCGAGAGCATCATTTATG 101
QY 61 AATATGAAAGGAGACCGAGCTGGGCTGCCAGAGGCAAGAGGCGCTTAAGCTTTTCA 120
DB 102 AATATGAAAGGAGACCGAGCTGGGCTGCCAGAGGCAAGAGGCGCTTAAGCTTTTCA 161
QY 121 AGCTACAAACAACGCTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
DB 162 AGCTACAAACAACGCTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 221
QY 181 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAAGCCGAAAGCAAGTGGTGAT 240
DB 222 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAAGCCGAAAGCAAGTGGTGAT 281
QY 241 ATGCTGGGAGACTGGGAGAAATTAACAAGAGAGAAAGTCAATGATGAGGCTCAAG 300
DB 282 ATGCTGGGAGACTGGGAGAAATTAACAAGAGAGAAAGTCAATGATGAGGCTCAAG 341
QY 301 GGAATCCCATGAACTCCGGGGCCGAGTGTGCTCACTCTCTGAACATTTAGAAATG 360
DB 342 GGAATCCCATGAACTCCGGGGCCGAGTGTGCTCACTCTCTGAACATTTAGAAATG 401
QY 361 AAGTTGAAAAACCCGGAAGATACAGATCATGAGAGAGAGGCAAGAGTCAATGAG 420
DB 402 AAGTTGAAAAACCCGGAAGATACAGATCATGAGAGAGAGGCAAGAGTCAATGAG 461
QY 421 CACATCAAGGATTCGACCGGAGCTTAACCGGAGCAATTAAGAGCAATATTTCTCAG 480
DB 462 CACATCAAGGATTCGACCGGAGCTTAACCGGAGCAATTAAGAGCAATATTTCTCAG 521
QY 481 GATCGATACGAAACCAAGCAGCGGAACTACTCCATCTCTGGCATATGAGAGTAT 540
DB 522 GATCGATACGAAACCAAGCAGCGGAACTACTCCATCTCTGGCATATGAGAGTAT 581
QY 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATGCGCGCTTGTCTCTCTAT 600
DB 582 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATGCGCGCTTGTCTCTCTAT 641
QY 601 CTTCCGAGAGAGATGCAATTTGGGAGCTGTGCACTGTGCGCAGTGAAGGCACTCC 660
DB 642 CTTCCGAGAGAGATGCAATTTGGGAGCTGTGCACTGTGCGCAGTGAAGGCACTCC 701
QY 661 CTGACGAGATTTCAACAGCCAAATGGCGGAAACGCTCAAGGCGCTCCAGACCAAGAG 720
DB 702 CTGACGAGATTTCAACAGCCAAATGGCGGAAACGCTCAAGGCGCTCCAGACCAAGAG 761
QY 721 CATGTGTAGCAACGTCACAAACCAAGCATATGGGCAATCAAGCAAGAAAGATCTATGT 780
DB 762 CATGTGTAGCAACGTCACAAACCAAGCATATGGGCAATCAAGCAAGAAAGATCTATGT 821
QY 781 GGGCAGTGTTCCTGTAAGCTGCTCATCCGAAATTTGATTGACGGGATCTTCTCGGG 840
DB 822 GGGCAGTGTTCCTGTAAGCTGCTCATCCGAAATTTGATTGACGGGATCTTCTCGGG 881
QY 841 CTACCCCTGCGCTGTGAGCGTGTATCTGTGAGAGGCAACAGGCTGTGATCCGATA 900

DB 882 CTCACCCCTGGCGCTGTGGACGCTGTATCTGTGAGAGGGAACAGGCGTTGATCCGATA 941
QY 901 ACAAGATGCGCTTTTAAGTTACAGAAAGCGCTTACAGAGCTGCAGGTGGCCG 960
DB 942 ACAAGATGCGCTTTTAAGTTACAGAAAGCGCTTACAGAGCTGCAGGTGGCCG 1001
QY 961 TGGGCAAGTTTGGCAACCGGTTGCTGATACCTGGGCGAGGAGTGAAGCACTGTGCTC 1020
DB 1002 TGGGCAAGTTTGGCAACCGGTTGCTGATACCTGGGCGAGGATGAAGCACTGTGCTC 1061
QY 1021 AAGCATCTTAAGGCGCTTATGAAGAACTAACAGAAACAGAGGAGACCTGCACCCCA 1080
DB 1062 AAGCATCTTAAGGCGCTTATGAAGAACTAACAGAAACAGAGGAGACCTGCACCCCA 1121
QY 1081 GCCAAACCGAGCAAGGCTGTGGCATTCAGGCGCTGTGCGGCTTCAAGTGGCGGAG 1140
DB 1122 GCCAAACCGAGCAAGGCTGTGGCATTCAGGCGCTGTGCGGCTTCAAGTGGCGGAG 1181
QY 1141 ACCCTTGAAGGAGGAGCAAGGAGCGGCTCCAGGCGCCACAGCCGCGTTCCCGCGGCC 1200
DB 1182 ACCCTTGAAGGAGGAGCAAGGAGCGGCTCCAGGCGCCACAGCCGCGTTCCCGCGGCC 1241
QY 1201 ATTGTGCACTTCCCGCACGAGGACCTGCTTCCACACCTGCTGCTGGTGGGCT 1260
DB 1242 ATTGTGCACTTCCCGCACGAGGACCTGCTTCCACACCTGCTGCTGGTGGGCT 1301
QY 1261 GTCCGGAAGACACTTACCTGTGGCACTCAGAGTGTGCCAGCCCGGCTGCTCAG 1320
DB 1302 GTCCGGAAGACACTTACCTGTGGCACTCAGAGTGTGCCAGCCCGGCTGCTCAG 1361
QY 1321 GGAAGACCTCAGGCTTCCGAGATTCGAGAGTGAATCCATGCCCCGCTCCCAAG 1380
DB 1362 GGAAGACCTCAGGCTTCCGAGATTCGAGAGTGAATCCATGCCCCGCTCCCAAG 1421
QY 1381 GACCTGAGAGTGAAGGCGCTTGTGCTCCGCAATTAATTAATTAATTAATTAATTA 1440
DB 1422 GACCTGAGAGTGAAGGCGCTTGTGCTCCGCAATTAATTAATTAATTAATTAATTA 1481
QY 1441 CGTGCCATATCCAGAGGAGCAAGCTGAGCGCTTGTGAGAGTGAATCCATGCCCCGCTCCCAAG 1500
DB 1482 CGTGCCATATCCAGAGGAGCAAGCTGAGCGCTTGTGAGAGTGAATCCATGCCCCGCTCCCAAG 1541
QY 1501 CGGCTGAGATCGGCTTGTGCTCAGACCACTGATTCGACAGGCAACCCCTTCA 1560
DB 1542 CGGCTGAGATCGGCTTGTGCTCAGACCACTGATTCGACAGGCAACCCCTTCA 1601
QY 1561 GCTAAGGAGCAAGCGGCTGTCTCCCACTCAGGCGCTTGTGCTGAGAGTGAATCCATGCCCCGCTCCCAAG 1620
DB 1602 GCTAAGGAGCAAGCGGCTGTCTCCCACTCAGGCGCTTGTGCTGAGAGTGAATCCATGCCCCGCTCCCAAG 1661
QY 1621 GAAAGTTCAGTTCCTTCCAGGCTTGAAGCAATGAGGCAAGGCTCATGAGTGAAT 1680
DB 1662 GAAAGTTCAGTTCCTTCCAGGCTTGAAGCAATGAGGCAAGGCTCATGAGTGAAT 1721
QY 1681 AATTTCCCTAAGCTTAAACCAAGCAAG-TTCCGATCTCTGTTTATTTTGTGTTAA 1740
DB 1722 AATTTCCCTAAGCTTAAACCAAGCAAG-TTCCGATCTCTGTTTATTTTGTGTTAA 1780
QY 1741 CTTATGAAAAATGATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800
DB 1781 CTTATGAAAAATGATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1840
QY 1801 AGACCCAGATCAAAAGCAACATGCCAGCGCTTCCAGACACCCCAAGCCCAAGCAG 1860
DB 1841 AGACCCAGATCAAAAGCAACATGCCAGCGCTTCCAGACACCCCAAGCCCAAGCAG 1900
QY 1861 CATGCTTGAATTTGAGAGCAACCGTGAAGCTGCTTGTATCTTAACTATGAAGAG 1920
DB 1901 CATGCTTGAATTTGAGAGCAACCGTGAAGCTGCTTGTATCTTAACTATGAAGAG 1960
QY 1921 ATAACTACCTTCAAGTTTGAATAATGTTTC 1953
DB 1961 ATAACTACCTTCAAGTTTGAATAATGTTTC 1993

Db 1729 GAAGTTCAGTTCCTCCAGGCTTCTAGAGCATCTGGCCAGGCTCATGGCTGAT 1788
 Qy 1681 AATTCCTTAGAGCTTAACAACCAAGCAAGCTTCCTCTCTTTATTTTGGTTAA 1740
 Db 1789 AATTCCTTAGAGCTTAACAACCAAGCAAGCTTCCTCTCTTTATTTTGGTTAA 1848
 Qy 1741 CTATGAAATGATTTAAGAAAGAGTGCAGCTCGAGAGATTCAGAGATGAAACACACC 1800
 Db 1849 CTATGAAATGATTTAAGAAAGAGTGCAGCTCGAGAGATTCAGAGATGAAACACACC 1908
 Qy 1801 AGACCCAGATCAACAAGCCATGCTCCAGCCCTCCAGCAACCCCAAGCCCAAGAC 1860
 Db 1909 AGACCCAGATCAACAAGCCATGCTCCAGCCCTCCAGCAACCCCAAGCCCAAGAC 1968
 Qy 1861 CATCTTTCTGAATTTCTGACGACACCGTGAAGCTCTCTTTTACTTTAACTCATGAAAG 1920
 Db 1969 CATCTTTCTGAATTTCTGACGACACCGTGAAGCTCTCTTTTACTTTAACTCATGAAAG 2028
 Qy 1921 ATACTACCTTCACGTTTGAATTAATGTTCTCTGTTGAATG 1964
 Db 2029 ATACTACCTTCACGTTTGAATTAATGTTCTCTGTTGAATG 2072

RESULT 4

US-10-302-172-691
 ; Sequence 691, Application US/10302172
 ; Publication No. US20040053250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 803_1CNCB
 ; CURRENT APPLICATION NUMBER: US/10/302,172
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/225,251
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: PCT US02/05095
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 09/799,451
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 950
 ; SOFTWARE: PL_FL_genes Version 2.0
 ; SEQ ID NO 691
 ; LENGTH: 2072
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (163)..(1743)
 ; US-10-302-172-691

Query Match 92.3%; Score 1812.4; DB 7; Length 2072;
 Best Local Similarity 96.3%; Pred. No. 0;
 Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

Qy 1 ATGACGTGTGATGAGTCCGCGGACGTTGTGGGCAACAAGGCGAGAGCATCATTTAG 60
 Db 163 ATGACGTGTGATGAGTCCGCGGACGTTGTGGGCAACAAGGCGAGAGCATCATTTAG 222
 Qy 61 AAATAGAAAGGAGACACGAGCTGCGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCA 120
 Db 223 AAATAGAAAGGAGACACGAGCTGCGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCA 282
 Qy 121 AGCTACAAACAACCTGATCATTTTGGGATTTGTATGATGAGAGAGCTGCTCTCTG 180
 Db 283 AGCTACAAACAACCTGATCATTTTGGGATTTGTATGATGAGAGAGCTGCTCTCTG 342
 Qy 181 ACTGCGGAGGAGGAGCAAAATTCGCGGAGATCAAGCCGAAAGCAAGTGGTGAAT 240
 Db 343 ACTGCGGAGGAGGAGCAAAATTCGCGGAGATCAAGCCGAAAGCAAGTGGTGAAT 402

Qy 241 ATCTGGAGAGTCTGGGAAATATCAAAAGCAGCAAGAGCTCATAGTCAGGCTCAAG 300
 Db 403 ATCTGGAGAGTCTGGGAAATATCAAAAGCAGCAAGAGCTCATAGTCAGGCTCAAG 462
 Qy 301 GGAATGCCATGAAATCCGCGGCTCCGATGTGTGCTCTCTCTGAAATTTGAGAAATG 360
 Db 463 GGAATGCCATGAAATCCGCGGCTCCGATGTGTGCTCTCTCTGAAATTTGAGAAATG 522
 Qy 361 AAGTTGAAATCCCGGAAATATCAATCATGAGAGAGGCAAGAGCTCATCTGAG 420
 Db 523 AAGTTGAAATCCCGGAAATATCAATCATGAGAGAGGCAAGAGCTCATCTGAG 582
 Qy 421 CATATCAGGAGATGACCGGAGCTAAGGAGGCAATTAAGAAAGATATATCTTCAAG 480
 Db 583 CATATCAGGAGATGACCGGAGCTAAGGAGGCAATTAAGAAAGATATATCTTCAAG 642
 Qy 481 GATGATAGGAAACCAAGAGCGGGAATCTATCAATCTCTCTGATATGAGAGAT 540
 Db 643 GATGATAGGAAACCAAGAGCGGGAATCTATCAATCTCTCTGATATGAGAGAT 702
 Qy 541 AACCCGAGGTGGCTATCTGACGAGACCTGAGCCATGCGCGCTTGTCTCTCTAT 600
 Db 703 AACCCGAGGTGGCTATCTGACGAGACCTGAGCCATGCGCGCTTGTCTCTCTAT 762
 Qy 601 CTTCCTGAGGAGATGATCTGAGGACCTGAGCTCTGAGGAGTGAAGGCACTCC 660
 Db 763 CTTCCTGAGGAGATGATCTGAGGACCTGAGCTCTGAGGAGTGAAGGCACTCC 822
 Qy 661 CTGACGAGATTTTCAAGCCCAATGCGGGAACCTGACAGGAGCTCCAAAGCAAGAG 720
 Db 823 CTGACGAGATTTTCAAGCCCAATGCGGGAACCTGACAGGAGCTCCAAAGCAAGAG 882
 Qy 721 CATGTGTAGCCACGTCAACAACCAAGCATGAGGATCAAGCAAAAGATCTATGT 780
 Db 883 CATGTGTAGCCACGTCAACAACCAAGCATGAGGATCAAGCAAAAGATCTATGT 942
 Qy 781 GGGCAATGTTCCCGGTAGGCTGCTCATCCGGAATTTGATGAGGAGATCTCTCGGG 840
 Db 924 -----GATCTCTCTCGGG 936
 Qy 841 CTCACCTCGGCTGTTGGAGCTGTATCTGTAGAGAGGAAACAGGCTGTGATCCGATA 900
 Db 937 CTCACCTCGGCTGTTGGAGCTGTATCTGTAGAGAGGAAACAGGCTGTGATCCGATA 996
 Qy 901 ACAAGATGCGCTTTAAGGTTCAAGAGAGCGCTCAAGAAAGCTCAAGTGTGCCCCG 960
 Db 997 ACAAGATGCGCTTTAAGGTTCAAGAGAGCGCTCAAGAAAGCTCAAGTGTGCCCCG 1056
 Qy 961 TGGGACGTTTGTGAAACCGGTTCTGTGATCTGTGGCCAGAGGATGAGACATGTGCTC 1020
 Db 1057 TGGGACGTTTGTGAAACCGGTTCTGTGATCTGTGGCCAGAGGATGAGACATGTGCTC 1116
 Qy 1021 AAGCATCTTAAAGGCTCTATGAAAGAACTAAACAAGAGAGGAGGAGCTGCCACCCCA 1080
 Db 1117 AAGCATCTTAAAGGCTCTATGAAAGAACTAAACAAGAGAGGAGGAGCTGCCACCCCA 1176
 Qy 1177 GCCAAACCCGAGCAAGAGGCTGTGCGGATCAAGGCTGTGCTGAGGCGGGAAG 1236
 Db 1141 ACCCTCTGAAAGGAGGAGAGAGGAGGCTTCCAGAGCCCAAGCCCGGTTCCCGAGGCC 1200
 Qy 1237 ACCCTCTGAAAGGAGGAGAGAGGAGGCTTCCAGAGCCCAAGCCCGGTTCCCGAGGCC 1296
 Db 1201 ATTGTGTAGCTTCCCGGCAAGGAGCACTGCTTTTCCACAACCTGTCTGTGGGAGCT 1260
 Qy 1297 ATTGTGTAGCTTCCCGGCAAGGAGCACTGCTTTTCCACAACCTGTCTGTGGGAGCT 1356
 Db 1261 GTCCGGAAGACACTTCTGTGGGCACTCAAGGTTGTGCCAGGCCGAGCTGCTCAG 1320
 Qy 1357 GTCCGGAAGACACTTCTGTGGGCACTCAAGGTTGTGCCAGGCCGAGCTGCTCAG 1416
 Db 1321 GGAAGACCTCAGGGTTCTGTGAGATTTCTGTGAGTCAATGCTCATGCTCCGAGAG 1380

Db	1417	GGAGGACCTCAGGGGTTCTTGAGAGTTCTCTGAGACTCTCATGCCCCGGCTTCCACG	1476
Qy	1381	GACCTGACGCTAGAGGGGCCCTTGGTTCCGCAATTATGATTTTCAGACAGAGCTGCGGTC	1440
Db	1477	GACCTGACGCTAGAGGGGCCCTTGGTTCCGCAATTATGATTTTCAGACAGAGCTGCGGTC	1536
Qy	1441	CGTGCATATTCGACGAGGAGACCACTATGGCCCCCTGCTGGGACGCTGAAACACCTGCGGAG	1500
Db	1537	CGTCCCATATCCGACGAGAGCACTATGGCCCCCTGCTGGGACGCTGAAACACCTGCGGAG	1596
Qy	1501	CGGGTGAATCGGCTTTTCCGCTGCACCCAGCAGCTGATTTCCGACCAAGGACACCCCTTCAGA	1560
Db	1597	CGGGTGAATCGGCTTTTCCGCTGCACCCAGCAGCTGATTTCCGACCAAGGACACCCCTTCAGA	1656
Qy	1561	GCTAGGGACCAACAGCCGCTGTGCTCCCACTCAAGGGCTTGCCTGTGGGACCTCCACTTGG	1620
Db	1657	GCTAGGGACCAACAGCCGCTGTGCTCCCACTCAAGGGCTTGCCTGTGGGACCTCCACTTGG	1716
Qy	1621	GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGGGACCAAGGCTATGAGCTGGAT	1680
Db	1717	GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGGGACCAAGGCTATGAGCTGGAT	1776
Qy	1681	AATTTCCCTAGGCTTAAACAACCAGAGGCTTGCCTCCCTGTTATTTTGGTTAA	1740
Db	1777	AATTTCCCTAGGCTTAAACAACCAGAGGCTTGCCTCCCTGTTATTTTGGTTAA	1836
Qy	1741	CTTATGAAAAATGTTATTAAGAAAAGTCACTCGAGAGAGATTCAAGATGAAACAACC	1800
Db	1837	CTTATGAAAAATGTTATTAAGAAAAGTCACTCGAGAGAGATTCAAGATGAAACAACC	1896
Qy	1801	AGACCCAGATCAAAAGCCAAACCATGCCAGCCCTCCAGACACCCCCAGCCCAACGAC	1860
Db	1897	AGACCCAGATCAAAAGCCAAACCATGCCAGCCCTCCAGACACCCCCAGCCCAACGAC	1956
Qy	1861	CATGCTTCTGAATTTCTGACGACACCGTAGGCTGCTTTGTACTTTAACTCAATGAAAG	1920
Db	1957	CATGCTTCTGAATTTCTGACGACACCGTAGGCTGCTTTGTACTTTAACTCAATGAAAG	2016
Qy	1921	ATAACTACTTCAAGTTTGAATAAATGTTTCTGTTGAAATG	1964
Db	2017	ATAACTACTTCAAGTTTGAATAAATGTTTCTGTTGAAATG	2060

```

RESULT 5
US-10-108-260A-676
; Sequence 676, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-108-260A-676

```

Query Match	92.0%	Score 1807.4	DB 6	Length 2547
Best Local Similarity	96.3%	Pred. No. 0		
Matches 1887	Conservative 0	Indels 6	Indels 66	Gaps 1
QY	1	ATGGAACGCTGTAAGCTCCGCGGACGTTGGTGGGCACAAGACCGAGACGATCATTTATG	60	
Db	755	ATGACGCTGTGAAGCTCGCGGACGTTGGTGGGCACAAGACCGAGACGATCATTTATG	814	
QY	61	AAATACCAAAAGGACACCCGAGCTGGGCTGCACAGGACAAGGGGCCCTAAACCTTTTTCGA	120	
Db	815	AAATACCAAAAGGACACCCGAGCTGGGCTGCACAGGACAAGGGGCCCTAAACCTTTTTCGA	874	

QY	121	AGCTACAA	CAACACGTCGATCATTTGGGGATTGTA	CATGAGACGGAGCTGCTCTCTG	180
Db	875	AGCTACAA	CAACACGTCGATCATTTGGGGATTGTA	CAATGAGACGGAGCTGCTCTCTG	934
QY	181	ACTGGCGGGAGGCGCAAGCAATTTGGCGGGAGATCAGCCGAAAGCAAGTGGGAGAT			240
Db	935	ACTGGCGGGAGGCGCAAGCAATTTGGCGGGAGATCAGCCGAAAGCAAGTGGGAGAT			994
QY	241	ATGCTGGGAGACTGGGAGAAATACAAAGCAGACAGAAAGCTCATAGATCAGCGTCAAG			300
Db	995	ATGCTGGGAGACTGGGAGAAATACAAAGCAGACAGAAAGCTCATAGATCAGCGTCAAG			1054
QY	301	GGAAATGCCATGAAACATCCGGGGCCCGATGTGTGTCAGTCTTCTTGAACTTGAAGAAATG			360
Db	1055	GGAAATGCCATGAAACATCCGGGGCCCGATGTGTGTCAGTCTTCTTGAACTTGAAGAAATG			1114
QY	361	AAAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGAAAGGCGAAGAGTCATCTGAG			420
Db	1115	AAAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGAAAGGCGAAGAGTCATCTGAG			1174
QY	421	CACATCCAGCGCATGACCGGGACGTAAAGCGGGACATTAAAGACATATATCTTCAGG			480
Db	1175	CACATCCAGCGCATGACCGGGACGTAAAGCGGGACATTAAAGAGATATATCTTCAGG			1234
QY	481	GATCGATA	CGGAACCAAGACGGGGAACTACTCCACATCTCTCTGSCATATGAGAGTAT		540
Db	1235	GATCGATACGGAACCAAGACGGGGAACTACTCCACATCTCTCTGSCATATGAGAGTAT			1294
QY	541	AAACCGGAGTGGGCTTACGTGACGGGACCTGAGCCCATTCGCGCTTGTTCCTCCCTAT			600
Db	1295	AAACCGGAGTGGGCTTACGTGACGGGACCTGAGCCCATTCGCGCTTGTTCCTCCCTAT			1354
QY	601	CTTCTCGAGAGATGATCATTTCTGGGCACTGGTGACGCTCTGCGCATGAGAGGCACTCC			660
Db	1355	CTTCTCGAGAGATGATCATTTCTGGGCACTGGTGACGCTCTGCGCATGAGAGGCACTCC			1414
QY	661	CTGACGGGATTTACAGCCCAATGCGGGACCGTCTCAGAGGCTCCAGAACCAACGAGAG			720
Db	1415	CTGACGGGATTTACAGCCCAATGCGGGACCGTCTCAGAGGCTCCAGAACCAACGAGAG			1474
QY	721	CATGTGTAGCCAGCTCACAAACCCAAAGACATGGGGCATCAGAGCAAGAAATCTATGT			780
Db	1475	CATGTGTAGCCAGCTCACAAACCCAAAGACATGGGGCATCAGAGCAAGAAATCTATGT			1515
QY	781	GGGCAGTGTCCCGGTAGCGTCAATCCGATATTGACGGGATCTCTCGGG			840
Db	1516	-----GATCTCTCTCGGG			1528
QY	841	CTCACCCCTGCGCTGTGGACGTGTATCTGTAGAAAGCGGAACAGCGTTGATCCGATAT			900
Db	1529	CTCACCCCTGCGCTGTGTGGACGTGTATCTGTAGAAAGCGGAACAGCGTTGATCCGATAT			1588
QY	901	ACAAGAAATCGCTTTAAAGTTTACAGAAAGCCCTTCAGAAACGTTCAGAGTGTGGCCCG			960
Db	1589	ACAAGAAATCGCTTTAAAGTTTACAGAAAGCCCTTCAGAAACGTTCAGAGTGTGGCCCG			1648
QY	961	TGGGACGTTTTTGCAACCGGTTGTTGATCTGGGACCAAGGATGAGGACACTGTGCTC			1020
Db	1649	TGGGACGTTTTTGCAACCGGTTGTTGATCTGGGACCAAGGATGAGGACACTGTGCTC			1708
QY	1021	AAGCATCTTAAGGCTCTTATGAAAGAACTTAACAAGAAACAGAGGGACCTGCAACCCCA			1080
Db	1709	AAGCATCTTAAGGCTCTTATGAAAGAACTTAACAAGAAACAGAGGGACCTGCAACCCCA			1768
QY	1081	GCCAAACCGAGCAAGGTCGTGCGCATTCAGGCTGTGCGGCTTCACTGTGCGGAAAG			1140
Db	1769	GCCAAACCGAGCAAGGTCGTGCGCATTCAGGCTGTGCGGCTTCACTGTGCGGAAAG			1828
QY	1141	ACCTCTGCAAGGGGAGCAGGAGGCGCTCCAGGCGCCACAGGCGCGTTCCGCGGCCC			1200
Db	1829	ACCTCTGCAAGGGGAGCAGGAGGCGCTCCAGGCGCCACAGGCGCGTTCCGCGGCCC			1888

```
QY 1201 ATTTGGTCAAGCTTCCCGGACGAGGACCTGTTCTTCCACACCTGCTCTGAGGGGCT 1260
DB 1889 ATTTGGTCAAGCTTCCCGGACGAGGACCTGTTCTTCCACACCTGCTCTGAGGGGCT 1348
QY 1261 GTCGGGAGAGACCTACCTGTTGGGCACTCAGGGTGTGCCAGCCCGGCTTGGCTCAG 1320
DB 1949 GTCGGGAGAGACCTACCTGTTGGGCACTCAGGGTGTGCCAGCCCGGCTTGGCTCAG 2008
QY 1321 GAGAGACCTCAGGGGTTCTTGGAGATTCTTGCATGTGAACTCCATGCCCTCCCAACG 1380
DB 2009 GAGAGACCTCAGGGGTTCTTGGAGATTCTTGCATGTGAACTCCATGCCCTCCCAACG 2068
QY 1381 GACCTGAGCTAGAGGGGCTTGGTTCCTGCACTTATGATTTTCAAGACAGCTGTGGTTC 1440
DB 2069 GACCTGAGCTAGAGGGGCTTGGTTCCTGCACTTATGATTTTCAAGACAGCTGTGGTTC 2128
QY 1441 CGTGCAATATCCAGAGAGACCAAGCTGAGCCCTGCTGAGGAGCTGAACCTGTGGAG 1500
DB 2129 CGTGCAATATCCAGAGAGACCAAGCTGAGCCCTGCTGAGGAGCTGAACCTGTGGAG 2188
QY 1501 CGGCTGAGATCGGCTTTCGTTGCAACCAAGCACTGATTTCCGACCAAGGACACCTTCAAG 1560
DB 2189 CGGCTGAGATCGGCTTTCGTTGCAACCAAGCACTGATTTCCGACCAAGGACACCTTCAAG 2248
QY 1561 GCTAGGAGACGAAAGCGGTGTGCTCCCACTCAGGGGCTTGGCTGAGGCTCCACTTG 1620
DB 2249 GCTAGGAGACGAAAGCGGTGTGCTCCCACTCAGGGGCTTGGCTGAGGCTCCACTTG 2308
QY 1621 GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGGGCAAGGCTCATGGCTGAT 1680
DB 2309 GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGGGCAAGGCTCATGGCTGAT 2368
QY 1681 AATTTCCCTTAGGCTTAAACAACCAAGCAAGCTTGGGCTCTCGTTTATTTTGGTTAAA 1740
DB 2369 AATTTCCCTTAGGCTTAAACAACCAAGCAAGCTTGGGCTCTCGTTTATTTTGGTTAAA 2428
QY 1741 CTTATGAAATGATTTAAAGAAAGTGAGAGCTGAGAGAGATTCAAGATGAAACACCC 1800
DB 2429 CTTATGAAATGATTTAAAGAAAGTGAGAGCTGAGAGAGATTCAAGATGAAACACCC 2488
QY 1801 AGACCCAGATCAAAAGCCAAACCATGAGCCCTCCAGACACCCCAAGCCCAAGAC 1860
DB 2489 AGACCCAGATCAAAAGCCAAACCATGAGCCCTCCAGACACCCCAAGCCCAAGAC 2548
QY 1861 CATGCTTCTGAATTTCTGACGACACCGTAGGCTGCTTGTACTTTAACTCATGGAAG 1920
DB 2549 CATGCTTCTGAATTTCTGACGACACCGTAGGCTGCTTGTACTTTAACTCATGGAAG 2608
QY 1921 ATTAACCTTCAAGTTTGAATTAATGTTCTCTGTG 1959
DB 2609 ATTAACCTTCAAGTTTGAATTAATGTTCTCTGTG 2647

RESULT 6
US-10-071-838-3
; Sequence 3, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2146
```

```
QY 1201 ATTTGGTCAAGCTTCCCGGACGAGGACCTGTTCTTCCACACCTGCTCTGAGGGGCT 1260
DB 1889 ATTTGGTCAAGCTTCCCGGACGAGGACCTGTTCTTCCACACCTGCTCTGAGGGGCT 1348
QY 1261 GTCGGGAGAGACCTACCTGTTGGGCACTCAGGGTGTGCCAGCCCGGCTTGGCTCAG 1320
DB 1949 GTCGGGAGAGACCTACCTGTTGGGCACTCAGGGTGTGCCAGCCCGGCTTGGCTCAG 2008
QY 1321 GAGAGACCTCAGGGGTTCTTGGAGATTCTTGCATGTGAACTCCATGCCCTCCCAACG 1380
DB 2009 GAGAGACCTCAGGGGTTCTTGGAGATTCTTGCATGTGAACTCCATGCCCTCCCAACG 2068
QY 1381 GACCTGAGCTAGAGGGGCTTGGTTCCTGCACTTATGATTTTCAAGACAGCTGTGGTTC 1440
DB 2069 GACCTGAGCTAGAGGGGCTTGGTTCCTGCACTTATGATTTTCAAGACAGCTGTGGTTC 2128
QY 1441 CGTGCAATATCCAGAGAGACCAAGCTGAGCCCTGCTGAGGAGCTGAACCTGTGGAG 1500
DB 2129 CGTGCAATATCCAGAGAGACCAAGCTGAGCCCTGCTGAGGAGCTGAACCTGTGGAG 2188
QY 1501 CGGCTGAGATCGGCTTTCGTTGCAACCAAGCACTGATTTCCGACCAAGGACACCTTCAAG 1560
DB 2189 CGGCTGAGATCGGCTTTCGTTGCAACCAAGCACTGATTTCCGACCAAGGACACCTTCAAG 2248
QY 1561 GCTAGGAGACGAAAGCGGTGTGCTCCCACTCAGGGGCTTGGCTGAGGCTCCACTTG 1620
DB 2249 GCTAGGAGACGAAAGCGGTGTGCTCCCACTCAGGGGCTTGGCTGAGGCTCCACTTG 2308
QY 1621 GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGGGCAAGGCTCATGGCTGAT 1680
DB 2309 GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGGGCAAGGCTCATGGCTGAT 2368
QY 1681 AATTTCCCTTAGGCTTAAACAACCAAGCAAGCTTGGGCTCTCGTTTATTTTGGTTAAA 1740
DB 2369 AATTTCCCTTAGGCTTAAACAACCAAGCAAGCTTGGGCTCTCGTTTATTTTGGTTAAA 2428
QY 1741 CTTATGAAATGATTTAAAGAAAGTGAGAGCTGAGAGAGATTCAAGATGAAACACCC 1800
DB 2429 CTTATGAAATGATTTAAAGAAAGTGAGAGCTGAGAGAGATTCAAGATGAAACACCC 2488
QY 1801 AGACCCAGATCAAAAGCCAAACCATGAGCCCTCCAGACACCCCAAGCCCAAGAC 1860
DB 2489 AGACCCAGATCAAAAGCCAAACCATGAGCCCTCCAGACACCCCAAGCCCAAGAC 2548
QY 1861 CATGCTTCTGAATTTCTGACGACACCGTAGGCTGCTTGTACTTTAACTCATGGAAG 1920
DB 2549 CATGCTTCTGAATTTCTGACGACACCGTAGGCTGCTTGTACTTTAACTCATGGAAG 2608
QY 1921 ATTAACCTTCAAGTTTGAATTAATGTTCTCTGTG 1959
DB 2609 ATTAACCTTCAAGTTTGAATTAATGTTCTCTGTG 2647

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
NAME/KEY: CDS
LOCATION: (1)..(1832)
US-10-071-838-3

Query Match 91.3%; Score 1794; DB 5; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1805; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 159 TGAAGGAGAGCTGCTCTCTGACCTGCGGAGGAGGAGCAATTCGCGGAGATTCAG 218
DB 342 TGAAGGAGAGCTGCTCTCTGACCTGCGGAGGAGGAGCAATTCGCGGAGATTCAG 401
QY 219 CCGAAAGACCAAGTGGTGGATATGCTGGGAGACTGGAGAAATACAAAGACAGAGAA 278
DB 402 CCGAAAGACCAAGTGGTGGT-GATATGCTGGGAGACTGGAGAAATACAAAGACAGAGAA 460
QY 279 GCTCATAGATCGAGCTCAAGGGAATGCCATGAACATCCGGGCTCCGATGTGTAGT 338
DB 461 GCTCATAGATCGAGCTCAAGGGAATGCCATGAACATCCGGGCTCCGATGTGTAGT 520
QY 339 CCTCTGAAATTTGAGGAAATGAAATGAAAAACCCCGGAAAGATACAGATCAATGAAGA 398
DB 521 CCTCTGAAATTTGAGGAAATGAAATGAAAAACCCCGGAAAGATACAGATCAATGAAGA 580
QY 399 GAAAGGCAAGAGTCACTGAGCACTCAAGGCACTCAAGGCACTCAAGGCACTCAAGGCACT 458
DB 581 GAAAGGCAAGAGTCACTGAGCACTCAAGGCACTCAAGGCACTCAAGGCACTCAAGGCACT 640
QY 459 AAGGAACATATATTTCTTCAAGGATGATACGGAACCAAGACAGGGAATCACTCAAT 518
DB 641 AAGGAACATATATTTCTTCAAGGATGATACGGAACCAAGACAGGGAATCACTCAAT 700
QY 519 CTTCTGAGATGAGAGATATTAACCGGAGGTGGCTCTGAGAGGACCTGAGCCACAT 578
DB 701 CTTCTGAGATGAGAGATATTAACCGGAGGTGGCTCTGAGAGGACCTGAGCCACAT 760
QY 579 CGCGGCTTGTCTCTCTCTATCTTCTGAGAGAGATGATTTCAAGCCCAATGGCGGACCTGCA 638
DB 761 CGCGGCTTGTCTCTCTCTATCTTCTGAGAGAGATGATTTCAAGCCCAATGGCGGACCTGCA 820
QY 639 GCTGAGCAATGAGGACCTCTCTGAGGGAATTTCAAGCCCAATGGCGGACCTGCA 698
DB 821 GCTGAGCAATGAGGACCTCTCTGAGGGAATTTCAAGCCCAATGGCGGACCTGCA 880
QY 699 GGGGCTTCAAGACCAAGAGACATGTGTAGCAAGTCAACCAAGCAATGGGGA 758
DB 881 GGGGCTTCAAGACCAAGAGACATGTGTAGCAAGTCAACCAAGCAATGGGGA 940
QY 759 TCAGGACAGAAAGATCTATGTGGGCAAGTTCCTCGTTAGGCTGCTCATCCGATATT 818
DB 941 TCAGGACAGAAAGATCTATGTGGGCAAGTTCCTCGTTAGGCTGCTCATCCGATATT 1000
QY 819 GATTAAGGATCTCTCTGAGGCTCAACCTGCGCTGAGGAGCTGTATCTGTAGAGG 878
DB 1001 GATTAAGGATCTCTCTGAGGCTCAACCTGCGCTGAGGAGCTGTATCTGTAGAGG 1060
QY 879 CGAACAGGCTTGAATGAGGCAATTAAGAAATGCTTTAAGGTTCAAGAGAGGCTTCA 938
DB 1061 CGAACAGGCTTGAATGAGGCAATTAAGAAATGCTTTAAGGTTCAAGAGAGGCTTCA 1120
QY 939 GAAAGGTCAGAGTGTGGGCGGAGGCACTTTTTCGAAACCGGTTGTTGATACCTGGG 998
DB 1121 GAAAGGTCAGAGTGTGGGCGGAGGCACTTTTTCGAAACCGGTTGTTGATACCTGGG 1180
QY 999 CAGGATGAGGACATGTGCTCAAGCATCTTAAAGGCTTATGAAGAAATCAAGAGAA 1058
DB 1181 CAGGATGAGGACATGTGCTCAAGCATCTTAAAGGCTTATGAAGAAATCAAGAGAA 1240
```

QY 1059 GCAGGGGACCTGCAACCCCGACCCGACCAAGGCTGTCGCGATCAGAGCCTGT 1118
DB 1241 GCAGGGGACCTGCAACCCCGACCCGACCAAGGCTGTCGCGATCAGAGCCTGT 1300
QY 1119 GCCGGCTTCACTGTCGGGAGAACCTTCTGCAAGGGGACAGGAGCCTCCAGGCC 1178
DB 1301 GCCGGCTTCACTGTCGGGAGAACCTTCTGCAAGGGGACAGGAGCCTCCAGGCC 1360
QY 1179 ACCAGCCCGGTTCCCGGGCCCACTTGGTTCAGCTTCCCGGACAGGAGCCTCGTCTTC 1238
DB 1361 ACCAGCCCGGTTCCCGGGCCCACTTGGTTCAGCTTCCCGGACAGGAGCCTCGTCTTC 1420
QY 1239 CACACCTGTCCTGTGGGAGCTGTCCGGGAAGACACTACCTGTGGCACTCAGGCTGT 1298
DB 1421 CACACCTGTCCTGTGGGAGCTGTCCGGGAAGACACTACCTGTGGCACTCAGGCTGT 1480
QY 1299 GCCCAGCCCGGCTGCTGCTGAGGAGACCTCAGGCTTCTGAGATTCCTGCACTGGA 1358
DB 1481 GCCCAGCCCGGCTGCTGCTGAGGAGACCTCAGGCTTCTGAGATTCCTGCACTGGA 1540
QY 1359 CTCCATGCCCCGCTCCCAACGAGACTGAGAGGAGCCTTGGTTCGCGATTAAGA 1418
DB 1541 CTCCATGCCCCGCTCCCAACGAGACTGAGAGGAGCCTTGGTTCGCGATTAAGA 1600
QY 1419 TTTCAGACAGAGCTGTGGGTCGCTGCCATATCCAGAGAGGACAGCTGGCCCCCTGCTG 1478
DB 1601 TTTCAGACAGAGCTGTGGGTCGCTGCCATATCCAGAGAGGACAGCTGGCCCCCTGCTG 1660
QY 1479 GCAGGCTGAAACCTCTGGAGGCGGGTGAGATCGGCTTCCCTGCACTCAGACTGATTC 1538
DB 1661 GCAGGCTGAAACCTCTGGAGGCGGGTGAGATCGGCTTCCCTGCACTCAGACTGATTC 1720
QY 1539 CGACCAAGGAGACCCCTTCAAGAGCTAGGAGGAGAAAGCGGTGCTCCCACTCAGAGGCC 1598
DB 1721 CGACCAAGGAGACCCCTTCAAGAGCTAGGAGGAGAAAGCGGTGCTCCCACTCAGAGGCC 1780
QY 1599 TTGCTCTGCGGCTCCCACTTGAAGTCTCAGTTCCTTCAGGCTTCTGAAGACT 1658
DB 1781 TTGCTCTGCGGCTCCCACTTGAAGTCTCAGTTCCTTCAGGCTTCTGAAGACT 1840
QY 1659 GGGCAGGAGCTCATGCTGAGTAATTTCCCTAGCTTAACAACCAAGCAAGCTTCCGCT 1718
DB 1841 GGGCAGGAGCTCATGCTGAGTAATTTCCCTAGCTTAACAACCAAGCAAGCTTCCGCT 1900
QY 1719 CCTGCTTTTATTTTGGTTAACTTATGAATAATGATTAAGAAGAGTGCAGTCAAG 1778
DB 1901 CCTGCTTTTATTTTGGTTAACTTATGAATAATGATTAAGAAGAGTGCAGTCAAG 1960
QY 1779 AGATTCAAGATGGAACAACACAGACCCAGATCAAAAGCCAACTGCGCAGCCCTC 1838
DB 1961 AGATTCAAGATGGAACAACACAGACCCAGATCAAAAGCCAACTGCGCAGCCCTC 2020
QY 1839 CCAGCACCCTCAGGCCCCAGCCATCGTTCTGAATTCAGAGACACCGTGAAGCTGCTT 1898
DB 2021 CCAGCACCCTCAGGCCCCAGCCATCGTTCTGAATTCAGAGACACCGTGAAGCTGCTT 2080
QY 1899 TGTACTTTAACTCATGGAAGATTAACCTTCAAGCTTTGAATAATGTTTCTGTT 1958
DB 2081 TGTACTTTAACTCATGGAAGATTAACCTTCAAGCTTTGAATAATGTTTCTGTT 2140
QY 1959 GAAATG 1964
DB 2141 GAAATG 2146

RESULT 7
US-10-071-838-5
; Sequence 5, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil

; APPLICANT: Peng, Yue
; APPLICANT: Tularix Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
; NAME/KEY: CDS
; LOCATION: (1) .. (1538)
US-10-071-838-5

Query Match 89.1%; Score 1750; DB 5; Length 1862;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1862; Conservative 0; Mismatches 0; Indels 102; Gaps 1;

QY 1 ATGAGAGTGTAGAGGTGCGGGCAGTTGTGGGCAACAAGCGAGAGACATCATTTATG 60
DB 1 ATGAGAGTGTAGAGGTGCGGGCAGTTGTGGGCAACAAGCGAGAGACATCATTTATG 60
QY 61 AAATACGAAAGGAGACCCGAGCTGGCTGCCAGAGACAGAGGCGCTTAAGCCTTTTCCA 120
DB 61 AAATACGAAAGGAGACCCGAGCTGGCTGCCAGAGACAGAGGCGCTTAAGCCTTTTCCA 120
QY 121 AGCTACAAACAACGCTGATCATTTGGGGATTTGATCATAGACGGAGCTGCCCTCTG 180
DB 121 AGCTACAAACAACGCTGATCATTTGGGGATTTGATCATAGACGGAGCTGCCCTCTG 180
QY 181 ACTGCGGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAGAAGCAAGTGGTGGAT 240
DB 181 ACTGCGGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAGAAGCAAGTGGTGGAT 240
QY 241 ATGCTGGAGACTGGGAAATTCAAAGAGCAAGAAAGCTCATAGTCAAGCTTCAAG 300
DB 241 ATGCTGGAGACTGGGAAATTCAAAGAGCAAGAAAGCTCATAGTCAAGCTTCAAG 300
QY 301 GGAATGCCATGAAACATCGGGGCGGATGTGCTCATGCTCTGAACTTGAAGAAATG 360
DB 301 GGAATGCCATGAAACATCGGGGCGGATGTGCTCATGCTCTGAACTTGAAGAAATG 360
QY 361 AAGTTGAAACCCCGGAAGATACAGATCATGAAGAGAGCAAGAGTCACTGAG 420
DB 361 AAGTTGAAACCCCGGAAGATACAGATCATGAAGAGAGCAAGAGTCACTGAG 420
QY 421 CACATCCAGGCAATCGACCGGAGCTGAAGGGACATTAAGAAAGCATATATCTTCAGG 480
DB 421 CACATCCAGGCAATCGACCGGAGCTGAAGGGACATTAAGAAAGCATATATCTTCAGG 480
QY 481 GATCGATAGGGAACCAAGCAGGGGAACTACTGCATGCTCTGGCATATGAGAGTAT 540
DB 481 GATCGATAGGGAACCAAGCAGGGGAACTACTGCATGCTCTGGCATATGAGAGTAT 540
QY 541 AACCCGAGGTGGCTACTGCAAGGACCTGAGCCATGCGCCCTTGTCTCTCTAT 600
DB 541 AACCCGAGGTGGCTACTGCAAGGACCTGAGCCATGCGCCCTTGTCTCTCTAT 600
QY 601 CTTCCTGAGGAGATGATTTCTGGGCACTGTGTCAGCTGTGCGCAGTGAAGCACTCC 660
DB 601 CTTCCTGAGGAGATGATTTCTGGGCACTGTGTCAGCTGTGCGCAGTGAAGCACTCC 660
QY 661 CTGCAAGGATTTTCAAGCCCAATATGCGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 720
DB 661 CTGCAAGGATTTTCAAGCCCAATATGCGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 720

QY 721 CATGGTAGCAGCAGTCAACAACCAAGACCATGGGAGTACGAGCAAGAAAGATCTATGT 780
DB 721 CATGGTAGCAGCAGTCAACAACCAAGACCATGGGAGTAC----- 761
QY 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATGTGATGACGGGATCTCTTCGGG 840
DB 762 ----- 761
QY 841 CTCACCCCTGCGCTGTGGGACGTGTATCTGTGTAGAAAGCGAACAAGGCTTGATGCGGATA 900
DB 762 -----GTATCTGTGTAGAAAGCGAACAAGGCTTGATGCGGATA 798
QY 901 ACAAGATCGCCTTTAAGTTTCAGCAGAAACGCTCCAGAAAGCCTCCAGGTGTGCGCG 960
DB 799 ACAAGATCGCCTTTAAGTTTCAGCAGAAACGCTCCAGAAAGCCTCCAGGTGTGCGCG 858
QY 961 TGGGCAAGTTTTCAGACCGGTTCTGTGATACCTGGGCAAGGATGAGACACTGTGCTC 1020
DB 859 TGGGCAAGTTTTCAGACCGGTTCTGTGATACCTGGGCAAGGATGAGACACTGTGCTC 918
QY 1021 AAGCATTTAGGGCCTCTATGAAAGAACTAAACAAGAAAGAGGAGGAGCCTGCCACCCCA 1080
DB 919 AAGCATTTAGGGCCTCTATGAAAGAACTAAACAAGAAAGAGGAGGAGCCTGCCACCCCA 978
QY 1081 GCCAAACCCGAGCAAGGAGTGTGCGGATCCAGGCGCTGTGCGGCTTCACTGGGCGGAG 1140
DB 979 GCCAAACCCGAGCAAGGAGTGTGCGGATCCAGGCGCTGTGCGGCTTCACTGGGCGGAG 1038
QY 1141 ACCCTGTGCAAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1039 ACCCTGTGCAAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098
QY 1201 AATTGTGAGCTTCCCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1099 AATTGTGAGCTTCCCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1158
QY 1261 GTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1159 GTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1218
QY 1321 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1219 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1278
QY 1381 GACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1279 GACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1338
QY 1441 CGTGCATATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1339 CGTGCATATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1398
QY 1501 CGGAGTGAAGTCCGCTTTCGTCGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1399 CGGAGTGAAGTCCGCTTTCGTCGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1458
QY 1561 GCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1459 GCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1518
QY 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTGAAAGCATCTGGGAGGAGGAGGAGGAGGAG 1680
DB 1519 GAAAGTTCTCAGTTCCTCCAGGCTTCTGAAAGCATCTGGGAGGAGGAGGAGGAGGAG 1578
QY 1681 AATTTCCTAGGCTTAAACAACCAAGCAAGCTTCCGCTCTCTTTTATTTTGTATAA 1740
DB 1579 AATTTCCTAGGCTTAAACAACCAAGCAAGCTTCCGCTCTCTTTTATTTTGTATAA 1638
QY 1741 CTTATGAAAGATGTTTAAAGAAAGAGTGCAGCTGAGAGAGATTCAAGATGAGAACACC 1800
DB 1639 CTTATGAAAGATGTTTAAAGAAAGAGTGCAGCTGAGAGAGATTCAAGATGAGAACACC 1698
QY 1801 AGAGCCAGATCAAGAGCAACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860

DB 1699 AGAGCCAGATCAAGAGCAACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1758
QY 1861 CATGTTCTGATTTCTGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
DB 1759 CATGTTCTGATTTCTGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1818
QY 1921 ATAACTACCTTCAGCTTTCAGAAATGTTTCTCTGTAATG 1964
DB 1819 ATAACTACCTTCAGCTTTCAGAAATGTTTCTCTGTAATG 1862

RESULT 8
US-10-094-466-61
Sequence 61, Application US/10094466
Publication No. US2003020363A1
GENERAL INFORMATION:
APPLICANT: Spvtek et al.
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 21402-290D
CURRENT APPLICATION NUMBER: US/10/094,466
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/288,148
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/338,375
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 60/275,579
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/335,302
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 60/275,601
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/276,000
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/277,338
PRIOR FILING DATE: 2001-03-20
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patin 2.1
SEQ ID NO 61
LENGTH: 1752
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (31) .. (1678)
US-10-094-466-61

Query Match 83.4%; Score 1638.8; DB 6; Length 1752;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 ATGAGAGGTGAGGTGCGGGCAGTTGTGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
DB 31 ATGAGAGGTGAGGTGCGGGTGTGTGTGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 90
QY 61 AATAAGAAAGGAGCAGCAGCTGGGCTGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 91 AATAAGAAAGGAGCAGCAGCTGGGCTGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 150
QY 121 AGCTACAAACAAGTGTATTTGGGAGTTGTATGATGAGAGCGAGAGTGCCTCTCTG 180
DB 151 AGCTACAAACAAGTGTATTTGGGAGTTGTATGATGAGAGCGAGAGTGCCTCTCTG 210
QY 181 ACTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240

Db 211 ACTCGCGGGAGTGAAGCAATTCCGGGGAGATCAGCCGAAAGCAATGGGTGAAA 270
Qy 241 ATGCTGGGAGCTGGGAGAAATACAAAAGCAGAGAAAGCTCATAGATCGAGCTCAAG 300
Db 271 ATGCTGGGAGAAATGGGACCTACAAAAGCAGAGAAAGCTCATAGATCGAGCTCAAG 330
Qy 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTCAAGTCTCTGAACATTTAGAGAAATG 360
Db 331 GGAATTCATGAACATCCGGGGCCCGATGTGTCAAGTCTCTGAACATTTAGAGAAATC 390
Qy 361 AAGTTGAAAAAATCCCGGAAAGATACAGATCATGAAGAGAGAGGCAAGGTCATCTGAG 420
Db 391 AAGTTGAAAAAATCCCGGAAAGATACAGATCATGAAGAGAGAGGCAAGGTCATCTGAA 450
Qy 421 CACATCCAGCGCATCGACCGGGAGCTGTAAGCGGGACATTAAGAAAGCATATTTCTCAG 480
Db 451 CACATCCAGCGCATCGAGACGTGAGCGGACATTAAGAGAGGCAATATTTCTCAGG 510
Qy 481 GATGATACGGAAACCAAGCAGCGGAACTACTCCATCTCTCGGCAATATGAGAGAT 540
Db 511 GATGATACGGAAACCAAGCAGCGGAACTACTCTCTCGGCAATATGAGAGAT 570
Qy 541 AACCCGAGGTGGCTTACTGAGGAGCCTGAGCCACATCGCCCTTGTCTCTCTAT 600
Db 571 AACCCGAGGTGGCTTACTGAGGAGCCTGAGCCACATCGCCCTTGTCTCTCTAT 630
Qy 601 CTTCCTGAGAGAGATGATCTTGTGGGCACTGTGACAGCTGTGGCAGTGAAGGACATCC 660
Db 631 CTTCCTGAGAGAGATGATCTTGTGGGCACTGTGACAGCTGTGGCAGTGAAGGACATCC 690
Qy 661 CTGCAAGGATTTCAAGCCCAATGAGCGGAGCCGTCCAGGGCTCCAAAGCAACAGAG 720
Db 691 CTGCAAGGATTTCAAGCCCAATGAGCGGAGCCGTCCAGGGCTCCAAAGCAACAGAG 750
Qy 721 CATGTGTAGCCAGCTCAACCAAGACATGAGGATGAGGCAAGAAAGATCTATGT 780
Db 751 CATGTGTAGCCAGCTCAACCAAGACATGAGGATGAGGCAAGAAAGATCTATGT 810
Qy 781 GGGAGAGTTCCTCCGTGAGGCTGCTCATCCGGGATTTGAGCGGATCTCTCGGG 840
Db 811 GGGAGAGTTCCTCTTAAGGCTGCTCATCCGAGATTTGAGCGGATCTCTCGGG 870
Qy 841 CTCAACCTGAGCTGTGGGAGCTGTATCTGTGAGAGGCGAAAGCGCTTATGCCATA 900
Db 871 CTCAACCTGAGCTGTGGGAGCTGTATCTGTGAGAGGCGAAAGCGCTTATGCCATA 930
Qy 901 ACAAGATCCGCTTTAAGGTTCAAGCAAGAGCGCTCACAGAGACGTCCAGGTGTGCCG 960
Db 931 ACAAGATCCGCTTTAAGGTTCAAGTACAGCGCTCACAGAGACGTCCAGGTGTGCCG 990
Qy 961 TGGGCAAGTTTTCGAAACCGGTTGTTGATACCTGGGCAAGGAGTGAAGAACTGTCTC 1020
Db 991 TGGGCAAGTTTTCGAAACCGGTTGTTGATACCTGGGCAAGGAGTGAAGAACTGTCTC 1050
Qy 1021 AAGCATCTTAAAGGCTCTAATGAAGAACTAACAAGAAAGCAGGGGAGCTTCCACCCCA 1080
Db 1051 AAGCATCTTAAAGGCTCTAATGAAGAACTAACAAGAAAGCAGGGGAGCTTCCACCCCA 1110
Qy 1081 GCCAAACCCGAGCAAGGTCGTGCGCATCAAGCCCTGTGCGGCTTCACTGCGGAGAG 1140
Db 1111 GCCAAACCCGAGCAAGGTCGTGCGCATCAAGCCCTGTGCGGCTTCACTGCGGAGAG 1170
Qy 1141 ACCCTTTCGAAAGGGGAGCAGGAGAGCCCTTCAGAGCCCAACAAGCCGTTCCCGCGGCC 1200
Db 1171 ACCCTTTCGAAAGGGGAGCAGGAGAGCCCTTCAGAGCCCAACAAGCCGTTCCCGCGGCC 1230
Qy 1201 ATTGTGTCAGCTTCCCGGCAAGGAGCCTGTCTTTCACACCTGTCTGTGTGGGCT 1260
Db 1231 ATTGTGTCAGCTTCCCGGCAAGGAGCCTGTCTTTCACACCTGTCTGTGTGGGCT 1290
Qy 1261 GTCCGGAGAGACATCACTCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGTGCTAG 1320

Db 1291 GTCCGGAGAGACATCACTCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGAG 1350
Qy 1321 GAGAGACTTCAGGGTTCTGGAAGATTCCTGACAGTGAATCTCAATGCCCGCTCCAAAG 1380
Db 1351 GAGAGACTTCAGGGTTCTGGAAGATTCCTGACAGTGAATCTCAATGCCCGCTCCAAAG 1410
Qy 1381 GACCTGACGTGAGAGGAGCCCTTGGTTCCGCAATTAATGATTTACAGACAGCTGTGGGTC 1440
Db 1411 GACCTGACGTGAGAGGAGCCCTTGGTTCCGCGATTAATGATTTACAGACAGCTGTGGGTC 1470
Qy 1441 GGTGCAATTCACAGAGAGACAGCTGAGCCCTTGTCTGAGAGCTGAACACCTGTGGAG 1500
Db 1471 GGTGCAATTCACAGAGAGACAGCTGAGCCCTTGTCTGAGAGCTGAACACCTGTGGAG 1530
Qy 1501 CGGGTGAATCGGCTTTCAGTGAAGCCAGGACATGATTCAGACAGAGGACCCCTTCAGA 1560
Db 1531 CGGGTGAATCGGCTTTCAGTGAAGCCAGGACATGATTCAGACAGAGGACCCCTTCAGA 1590
Qy 1561 GCTAGGAGCAAGCCGCTGTGCTCCACCTCAGGGCTTGTCTGTGCGGCTTCACTTG 1620
Db 1591 GCTAGGAGCAAGCCGCTGTGCTCCACCTCAGGAGCTTGTGCTGTGCGGCTTCACTTG 1650
Qy 1621 GAAAGTTCTCATGTTCCCTCCAGGCTTTCAGAGCATTTGGGCGCAGGCTCATGCTGGAT 1680
Db 1651 GAAAGTTCTCATGTTCCCTCCAGGCTTTCAGAGCATTTGGGCGCAGGCTCATGCTGGAT 1710
Qy 1681 AATTTCCCTAAGGCTTAACAACCAAGCAGCTTTCAGCTCTC 1722
Db 1711 AATTTCCCTAAGGCTTAACAACCAAGCAGCTTTCAGCTCTC 1752

RESULT 9
US-10-450-763-21335
: Sequence 21335, Application US/10450763
: Publication No. US20050196754A1
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc
: TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 790CIP3/US
: CURRENT APPLICATION NUMBER: US/10/450, 763
: PRIOR FILING DATE: 2003-06-11
: PRIOR APPLICATION NUMBER: PCT/US01/08631
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: 09/540,217
: PRIOR FILING DATE: 2000-03-31
: PRIOR APPLICATION NUMBER: 09/649,167
: PRIOR FILING DATE: 2000-08-23
: NUMBER OF SEQ ID NOS: 60736
: SOFTWARE: Custom
: SEQ ID NO 21335
: LENGTH: 7856
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SIMILAR
: LOCATION: (1697)..(4054)
: OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number
: OTHER INFORMATION: X63546, Smith-Waterman Score=4233.
US-10-450-763-21335

Query Match 62.1%; Score 1220; DB 9; Length 7856;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
Qy 1 ATGACGCTGTGATAGCTCGCGGAGCATGTTGTGGGCAACAAGGAGAGCATCATATATG 60
Db 1697 ATGACATGTGTAGAGATGCAATAGTTTCAGGCAACAGAGCGGAAAGCATATATATG 1756
Qy 61 AAATGCAAAAAGGAGACACGAGCTGGGCTGCAAGAGACAAAGGGGCTTAAGCTTTTGA 120
Db 1757 AAGTATGCAAGGAGACACGAGCTGGGCTGCAAGAGACAAAGGGGCTTAAGCTTTTGA 1816
Qy 121 AGTACAACAACAACGTGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 180

D	b	1817	ATCAACAGAGCA---TTGATGCTTTTGGCAATTTTGCATGTAGACGGAAGCTGCTCTCTGTG	1873
O	y	181	ACTGCGCGGAGGCGCAAGCAAAATTCGCGCGGAGATCAAGCCGAAAGCAAGTGGGTGAT	240
D	b	1874	ACTGCACGGGAGGCGAAGAAATTTCCGCGCGGAGATGACACGAAACGACAAAGTGGATGAA	1933
O	y	241	ATGCTGGGAGACTGGGAGAAATTCAAAAAGCAGCAAAAAGCTCATGTAGTCAGCTTACAG	300
D	b	1934	ATGCTGGGAGAAATGGGAGCAATTAAGCAAGTACCAATCTATGATCGAGTGTACAAG	1993
O	y	301	GGAAATGCCATGAATCATCCGGGGCCCGATGTGGTCAAGCTTCCTGTAACATTTAGGGAATG	360
D	b	1994	GGAAATTTCCATGAACATCCGGGGCCCGGTGTGGTCAAGTCTCTTAACATTCAGGAATTC	2053
O	y	361	AAATTGAAAAAATCCCGGAAGATACCAATCATGAAGAAAGGCGAAGGTCAATCTGAG	420
D	b	2054	AAATTGAAAAAATCCCGGAAGATACCAATCATGAAGAAAGGCGAAGGTCAATCTGAA	2113
O	y	421	CACATCCAGCGCATGACCCGGGAGCTTAAGCGGAGACATTAAGAGACATATATCTTCAG	480
D	b	2114	CACATCCACCAATGACCTGGAAGCTGAAGACGACTCTCCGAAACATATCTCTTTAGG	2173
O	y	481	GATCGATACGGAACCAAGCAGCGGGAACTAATCCCATCTCTCGGCAATATGAGGAGAT	540
D	b	2174	GATCGATATGAGACCAAGCAGGGAACTAATCTAATCTCTCGGCAATATGAGGAGAT	2233
O	y	541	AAACCGAAGTGGGCTACTGCAAGGACCTGAGCCATCGCCGCTTGTCTCTCTAT	600
D	b	2234	AAACCGAAGTGGGCTACTGCAAGGACCTGAGCCATCAACGCTTGTCTCTCTAT	2293
O	y	601	CTTCTTGAAGAGATGCAATTTGGGCACTGGTCAAGCTCTGGCCAGTGAAGCACTCC	660
D	b	2294	CTGCTTGAAGAGAGCAATTTGGGCACTGGTCAAGCTCTGGCCAGTGAAGCACTCC	2353
O	y	661	CTGCAAGGATTTCAAGCCCAAATGGGAGGAAACCGTCAAGGGGCTTCAAGCAACAGAG	720
D	b	2354	CTGCAAGGATTTCAAGCCCAAATGGGAGGAAACCGTCAAGGGGCTTCAAGCAACAGAG	2413
O	y	721	CATGTGTAGCAAGCTCAAAACCCAGACCAATGGGCACTCAGACAAAGAAAGATCTATGT	780
D	b	2414	CATGTGTAGCAAGCTCAAAACCCAGACCAATGGGCACTCAGACAAAGAAAGATCTATGT	2473
O	y	781	GGGCAAGTTCCTCCGTTAGGCTGCTCATCCGATATGATTTGACGGGATCTCTCTCGG	840
D	b	2474	GGGCAAGTTCCTCCGTTAGGCTGCTCATCCGATATGATTTGACGGGATCTCTCTCGG	2533
O	y	841	CTCACCCCTGGGCTGTGGAGGTGTATCTGGTAAAGGCGAAACAGGCTTGAATGCCATA	900
D	b	2534	CTCACCCCTGGGCTGTGGAGGTGTATCTGGTAAAGGCGAAACAGGCTTGAATGCCATA	2593
O	y	901	ACAAGAAATCGCTTTTAAGGTTCAAGCAGAAAGGCTCAAGAAAGTCAAGGTGTGGCCG	960
D	b	2594	ACAAGAAATCGCTTTTAAGGTTCAAGCAGAAAGGCTCAAGAAAGTCAAGGTGTGGCCG	2653
O	y	961	TGGGACGTTTTCGCAACCGGTTCTGTGTATACCTGGGCAAGGATGAAGCACTGTGCTC	1020
D	b	2654	TGGGACGTTTTCGCAACCGGTTCTGTGTATACCTGGGCAAGGATGAAGCACTGTGCTC	2713
O	y	1021	AAGGATCTTAAAGGCTCTATGAAGAACTTAAGAAAGAGAGGGGACCTTGCCACCCCA	1080
D	b	2714	AAGGATCTTAAAGGCTCTCTACGAAAGAACTTAAGAAAGAGAGGGGACCTTGCCACCCCA	2773
O	y	1081	GCCAAACCGAGCAAGGGTCTGTCGCAATCAGAGCCTGACCGGCTTCAAGTGTGGGAG	1140
D	b	2774	GCCAAACCGAGCAAGGGTCTTGTGCAACCCAGGCTGTGTGCGCTTCAAGTGTGTGGAG	2833
O	y	1141	ACCCTCTGCAAGGGGAGACAGGACAGGACCTTCAGAGCCCAACAGCCGAGTTCGCGGCC	1200
D	b	2834	ACCCTCTGCAAGGGGATATAGGCAAGGACCTTCAGAGCCCAACAGCCGAGTTCGCGGCC	2893
O	y	1201	ATTGGTCAAGTTTCCCGCAACGGGCACTGTGTTCTTCAACACCTGTCTGTGTGGGCT	1260

QY	DB	Sequence	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	2894	ATTGTGTCAGTTTCCCCCGCATGGGACATCTCGTTTTCACGCGCCCTGTCTGGTGGGGCT	2953							
QY	1261	GTCCGGGAAGACACTTACCTCTGTGGGACCTCAGGGTGTGCCACCCCGGCTGGCTCAG	1320							
DB	2954	GTCCGGGAAGACACTTACCTCTGTGGGACCTCAGGGTGTGCCACCCCGGCTGGCTCAG	3013							
QY	1331	GGAGGACCTCAGGGTCTCTGGAGATTCTGTGAGTGGAACTTCATGCCCCGCTTCCACG	1380							
DB	3014	GGAGGACCTCAGGGTCTCTGGAGATTCTGTGAGTGGAACTTCATGCCCCGCTTCCACG	3073							
QY	1381	GACCTGGACGAGAGGGGCGCTGGTTCGGCATTTAGATTTCAGACAGAGCTGTGGGTC	1440							
DB	3074	GACCTGGATATAGGGGCGCTGGTTCGGCATTTAGATTTCAGACAGAGCTGTGGGTC	3133							
QY	1441	CGTGCCATATCCACGAGAGACCAAGCTGGCCCCCTGTGTGAGGCTGAACACCCCTGGCGAG	1500							
DB	3134	CGTGCCATATCCACGAGAGACCAAGCTGGCCCCCTGTGTGAGGCTGAACACCTGGCGAG	3193							

RESULT 10

US-10-450-763-24017

Sequence 24017, Application US/10450763

Publication No. US20050196754A1

GENERAL INFORMATION:

APPLICANT: Hyeq, Inc

TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 790CIP3/US

CURRENT APPLICATION NUMBER: US/10/450, 763

CURRENT FILING DATE: 2003-06-11

PRIOR APPLICATION NUMBER: PCT/US01/08631

PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 09/540, 217

PRIOR FILING DATE: 2000-03-31

PRIOR APPLICATION NUMBER: 09/649, 167

PRIOR FILING DATE: 2000-08-23

NUMBER OF SEQ ID NOS: 60736

SOFTWARE: Custom

SEQ ID NO 24017

LENGTH: 7856

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SIMILAR

LOCATION: (1697)..(4054)

OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number US-10-450-763-24017

Query Match

Best Local Similarity 88.8%; Pred. No. 0;

Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 1 ATGGAAGTGTATAGAGTTCGGGGCAATTTGTGGGCAAGAGTGGAGACATCATTAAG 60

DB 1697 ATGGAAGTGTATAGAGATGCAATGATTTTGCAGGCAAGAGCGGAAGACATCAATTAAG 1756

QY 61 AAATACGAAAAGGAGACACCGAGCTGGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCA 120

DB 1757 AAGTATGACAAAGGAGACACCGAGCTGGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCA 1816

QY 121 AGCTACAAACAACGCTGCATCTTTGGGGATTGTACATGAGACGAGAGTGCCTCTCTG 180

DB 1817 ATCAACAGCAGCA---TTGATCGTTTGGCATTTTGCATGAGACGAGAGTGCCTCTCTG 1873

QY 181 ACTGCGCGGGAGGGCGAAGCAAAATTCGGCGGGAGATATAGCCGAAAGGCAAGTGGTGGAT 240

DB 1874 ACTGCAACGGGAGGCGAAGAAATTCGGCGGGAGATATAGCAAGAACGCAAGTGGATGGAA 1933

QY 241 ATCTGTGGAGACTGGGAGAAATTCAAAAGCAGACGAAAGAACTATAGATGAGCTTACAG 300

DB 1934 ATCTGTGGAGAAATGGAGAGACATTTAGACACAGTAACTATAGATGAGATGTACAG 1993

QY 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGATCTTCTGGAACATTTGAGGAAATG 360

Db 1994 GGAATTCATGAAATCCGGGGCCCGGTGTGTCAAGTCTCTGAAATTCAGAAATC 2053
Qy 361 AAGTTGAAAAAATCCCGGAAGATACCAATCATGAAAGAAAGGCAAGATCATCTGAG 420
Db 2054 AAGTTGAAAAAATCCCGGAAGATACCAATCATGAAAGAAAGGCAAGATCATCTGAA 2113
Qy 421 CACATCCAGGCGATCGACCGGGAGGTAAAGCGGACATTAAGAAAGCATATATCTTCAG 480
Db 2114 CACATCCAGGCGATCGACCGGGAGGTAAAGCGGACATTAAGAAAGCATATATCTTCAG 2173
Qy 481 GATGATATCGGAACCAAGCAAGCGGGAACTACTCAATCTCTGCGCATATGAGAGAT 540
Db 2174 GATGATATCGGAACCAAGCAAGCGGGAACTACTCAATCTCTGCGCATATGAGAGAT 2233
Qy 541 AACCCGAGGTGGGCTACTGAGGAGACTGAGCCCAATCGCCGCTTGTCTCTCTAT 600
Db 2234 AACCCGAGGTGGGCTACTGAGGAGACTGAGCCCAATCGCCGCTTGTCTCTCTAT 2293
Qy 601 CTTCCTGAGGAGATGATCTGAGGCACTGGTGCAGCTGGGCAAGTGAAGGCACTCC 660
Db 2294 CTTCCTGAGGAGATGATCTGAGGCACTGGTGCAGCTGGGCAAGTGAAGGCACTCC 2353
Qy 661 CTGCAAGGATTTCAAGCCCAATGCGGGACCGTCCAGGGGCTCCAAAGCAAGAG 720
Db 2354 CTGCAAGGATTTCAAGCCCAATGCGGGACCGTCCAGGGGCTCCAAAGCAAGAG 2413
Qy 721 CATGTGTAGCGACGTCAACCAAGCAAGCGGCTCTGAGCAAGAAAGATCTATGT 780
Db 2414 CATGTGTAGCGACGTCAACCAAGCAAGCGGCTCTGAGCAAGAAAGATCTATGT 2473
Qy 781 GGGCAGTGTCCCGGTAGGCTGCTCATCCGGAATATGATGACGGGATCTCTCGGG 840
Db 2474 GGGCAGTGTCCCGGTAGGCTGCTCATCCGGAATATGATGACGGGATCTCTCGGG 2533
Qy 841 CTCAACCTGCGCTGTGAGGAGGTATCTGTGAGAAAGCAAGCGTGTATGATGAT 900
Db 2534 CTCAACCTGCGCTGTGAGGAGGTATCTGTGAGAAAGCAAGCGTGTATGATGAT 2593
Qy 901 ACAAGATTCGCTTTAAGTTACAGCAAGAGCGCTCAAGAAAGTCCAGGTGTGCGCG 960
Db 2594 ACAAGATTCGCTTTAAGTTACAGCAAGAGCGCTCAAGAAAGTCCAGGTGTGCGCG 2653
Qy 961 TGCGCAGGTTTTCACACCGGCTGTGATGATCTGAGGCAAGGATCACTGTGCTC 1020
Db 2654 TGCGCAGGTTTTCACACCGGCTGTGATGATCTGAGGCAAGGATCACTGTGCTC 2713
Qy 1021 AAGCATTTAGGAGCTTATGAAAGAACTAAAGAAAGCAAGGAGGAGCTGCCA 1080
Db 2714 AAGCATTTAGGAGCTTATGAAAGAACTAAAGAAAGCAAGGAGGAGCTGCCA 2773
Qy 1081 GCCAAACCGGAGAGGCTGTGCGCATCCAGGCTGTGCGGCTTCAAGTGTGAGAG 1140
Db 2774 GCCAAACCGGAGAGGCTGTGCGCATCCAGGCTGTGCGGCTTCAAGTGTGAGAG 2833
Qy 1141 ACCCTCTGCAAGGGGAGCAGGCAAGGCGCTCTCAAGGCCACAGGCCGAGTCC 1200
Db 2834 ACCCTCTGCAAGGGGATATGAGCAGGCGCTCTCAAGGCCACAGGCCGAGTCC 2893
Qy 1201 ATTGTGTCAGCTTCCCGGCAAGGAGCACTGTCTTTCACACCTGTGCTGTGAGGCT 1260
Db 2894 ATTGTGTCAGCTTCCCGGCAATGAGCACTGTCTTTCACAGCCCTGTGCTGTGAGGCT 2953
Qy 1261 GTCCGGGAGACACTCACTCTGTGAGCACTGAGGCTGTGAGGCGGCTGTGAG 1320
Db 2954 GTCCGGGAGACAGGTATCCCTGTGAGCACTGAGGCTGTGAGGCGGCTGTGAG 3013
Qy 1321 GGAAGACCTCAGGGGCTCTGAGATCTCTGAGTGAATCTCATGAGCCCGGCTCC 1380
Db 3014 GGAAGACCTCAGGGGCTCTGAGATCTCTGAGTGAATCTCATGAGCCCGGCTCC 3073
Qy 1381 GACCTGAGCGTGAAGGGGCTTGTGCTCGGCATTAATGATTTCAAGAGAGCTGTGAGT 1440

Db 3074 GACCTGATATTAAGGGGCGCTTGATTCCTCCCATTAATGATTTTGAACGAGCTGTGAGT 3133
Qy 1441 GGTGCAATATCCAGAGAGCAAGCTGGGCGCTGTGAGGCTGAACACCTGTGAGG 1500
Db 3134 GTGCAATATCCAGAGAGCAAGCTGGGCGCTTGATTCCTCCCATTAATGATTTTGAACGAGCTGTGAGT 3193

RESULT 11
US-09-962-436-562
; Sequence 562, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppel, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 562
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-562

Query Match 62.1%; Score 1220; DB 3; Length 7878;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy 1 ATGAGCTGTGATAGTCCGGGCGATGTTGATGGGCAACAGAGCAAGAGGATCATCTATG 60
Db 1697 ATGAGCAATGTGTAGAGATGCAATGTTTGCAGGCAAGAGCGGAAAGATCATCTATG 1756
Qy 61 AAATACGAAAGGAGACACGAGCTGGGCTGCAAGAGCAAGAGGCGCTTAAGCCTTTTGA 120
Db 1757 AAGTATGACAAAGGACACGAGCTGGGCTGCAAGAGCAAGAGGCGCTTAAGCCTTTTGA 1816
Qy 121 AGCTACAAACAACGCTCATTTTGGGATTTGTATGATGACGAGCTGCTCTCTG 180
Db 1817 ATCAAGCAGCA---TTGATCGTTTGGCATTTTGCATGACGAGCTGCTCTCTG 1873
Qy 181 ACTGCGGGAGGCGAAGCAATTCGCGGAGATGACCGGAAAGCAAGTGGTGTGAT 240
Db 1874 ACTGCAAGGAGGCGAAGAAATTCGCGGAGATGACGAAAGCAAGTGGTGTGAT 1933
Qy 241 ATGCTGGGAGACTGGGAGAAATCAAAAGCAGCAGAAAGCTCATAGATCGAGGTACA 300
Db 1934 ATGCTGGGAGATGGGAGAAATTAAGCAGATGACAAATCTCATAGATCGAGGTACA 1993
Qy 301 GGAATGCCATGAACATCCGGGCGGATGTGATCACTCTCTGAACATTTGAGAAATG 360
Db 1994 GGAATTCATGAACATCCGGGCGGATGTGATCACTCTCTGAACATTTGAGAAATC 2053
Qy 361 AAGTTGAAAAAATCCCGGAAGATACCAATCATGAAAGAAAGGCAAGATCATCTGAG 420
Db 2054 AAGTTGAAAAAATCCCGGAAGATACCAATCATGAAAGAAAGGCAAGATCATCTGAA 2113
Qy 421 CACATCCAGGCGATCGACCGGGAGGTAAAGCGGACATTAAGAAAGCATATATCTTCAG 480
Db 2114 CACATCCAGGCGATCGACCGGGAGGTAAAGCGGACATTAAGAAAGCATATATCTTCAG 2173
Qy 481 GATGATATCGGAACCAAGCAAGCGGGAACTACTCAATCTCTGCGCATATGAGAGAT 540
Db 2174 GATGATATCGGAACCAAGCAAGCGGGAACTACTCAATCTCTGCGCATATGAGAGAT 2233
Qy 541 AACCCGAGGTGGGCTTACTGAGGAGACTGAGCCCAATCGCCGCTTGTCTCTCTAT 600
Db 2234 AACCCGAGGTGGGCTTACTGAGGAGACTGAGCCCAATCGCCGCTTGTCTCTCTAT 2293

QY 601 CTTCTGAGAGAGATGATCTGCGGCACTGCTGAGCTGCTGCGCATGTGAGGCACTCC 660
|||
DB 2294 CTGCTGAGAGAGAGCATTTCTGGGCACTGTGTGAGCTGTGGCGCATGTGAGGCACTCC 2353
QY 661 CTGCAAGGATTTTCAACAGCCCAATGCGGGGACCGTCCAGGGGCTCCAAAGCAACAGGAG 720
|||
DB 2354 CTGCGAGGATTTCAACAGCCCAATGCGGGGACCGTCCAGGGGCTCCAAAGCAACAGGAG 2413
QY 721 CATGTGGTACGCAAGTCAACCAAGACATGAGGGGATGAGGACAAAGAAATCTATGT 780
|||
DB 2414 CATGTGGTACGCAAGTCAACCAAGACATGAGGGGATGAGGACAAAGAAATCTATGT 2473
QY 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTGACGGGATCTCTCGGG 840
|||
DB 2474 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTGACGGGATCTCTCGGG 2533
QY 841 CTCAACCTGCGCTGTGGGACGTGTATCTGCTAGAAAGGCAACAGCGTTGATGCGGATA 900
|||
DB 2534 CTCAACCTGCGCTGTGGGACGTGTATCTGCTAGAAAGGCAACAGCGTTGATGCGGATA 2593
QY 901 ACAAGAAATGCGCTTTAAGTTTCAAGCAAGAGGCGCTCCAGAAAGCTCAAGTGTGCGCG 960
|||
DB 2594 ACCAGCAATGTCTCTTAAGTTTCAAGCAAGAGGCGCTCCAGAAAGCTCAAGTGTGCGCG 2653
QY 961 TGGGCAAGTTTGGCAACCGGTTCTGTATACCTGGGCGCAGGATGAGGACACTGTGCTC 1020
|||
DB 2654 TGGGCAAGTTTGGGCAACCGGTTCTGTATACCTGGGCGCAGGATGAGGACACTGTGCTC 2713
QY 1021 AAGCATTTTGGGCTCTTATGAAAGAACTAAAGAAAGAGGGGGACCTGCCACCCCA 1080
|||
DB 2714 AAGCATTTTGGGCTCTTATGAAAGAACTAAAGAAAGAGGGGGACCTGCCACCCCA 2773
QY 1081 GCCAAACCGAGAGAGGGGTGTGGGCACTCAAGGCTGTGCGGCTTACAGTGTGGGAG 1140
|||
DB 2774 GCCAAACCGAGAGAGGGGTGTGGGCACTCAAGGCTGTGCGGCTTACAGTGTGGGAG 2833
QY 1141 ACCCTCTGCAAGGGGAGAGAGGAGGCGCTCTCAAGGCGCCACAGCGCGGCTCCGCGGCC 1200
|||
DB 2834 ACCCTCTGCAAGGGGATTAAGGAGGCGCTCTCAAGGCGCCACAGCGCGGCTCCGCGGCC 2893
QY 1201 ATTTGGTCAAGCTTCCCGCAACGGGCACTGTGTTCTTCAACACCTGTCTGTGGGGCT 1260
|||
DB 2894 ATTTGGTCAAGCTTCCCGCAACGGGCACTGTGTTCTTCAACACCTGTCTGTGGGGCT 2953
QY 1261 GTCGGGGAAGACACTACCTGCTGGGAGCTCAAGGGTGTGCCAGCGCGGCGCTGGCTCAG 1320
|||
DB 2954 GTCGGGGAAGACACTACCTGCTGGGAGCTCAAGGGTGTGCCAGCGCGGCGCTGGCTCAG 3013
QY 1321 GGAAGACCTCAAGGGTCTCTGAGATTCCTGCACTGAACTCAAGCGCGCTCCCAAG 1380
|||
DB 3014 GGAAGACCTCAAGGGTCTCTGAGATTCCTGCACTGAACTCAAGCGCGCTCCCAAG 3073
QY 1381 GACCTGAGACGTAGAGGGCGCTTGGTTCCGCAATTAATTAATTTCAAGACAGAGCTGTGGGTC 1440
|||
DB 3074 GACCTGAGATTAAGAGGGCGCTTGGTTCCGCAATTAATTAATTTCAAGACAGAGCTGTGGGTC 3133
QY 1441 CGTGCAATATCCAGAGAGAGCAAGCTGGCGCGCTGCGGAGAGGTGAACACCTGCGGAG 1500
|||
DB 3134 CGTGCAATATCCAGAGAGAGCAAGCTGGCGCGCTGCGGAGAGGTGAACACCTGCGGAGAG 3193

RESULT 13

US-10-956-157-1259
; Sequence 1259, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Myelch
; APPLICANT: Mounes, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04

QY 1000 NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1259
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1259
Query Match 62.1%; Score 1220; DB 9; Length 7878;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 1 ATGGAAGTGTAGAGGTGCGCGGAGTTGTGTGGGCAAAAGCGAGAGGACATCATTTATG 60
|||
DB 1697 ATGGAAGTGTAGAGGTGCGCGGAGTTGTGTGGGCAAAAGCGAGAGGACATCATTTATG 1756
QY 61 AAATACGAAAGGAGACACGAGCTGGGCTGGCCAGAGGACAAAGGGCTTAAGCTTTTCCA 120
|||
DB 1757 AAGTATGACAAAGGAGACACGAGCTGGGCTGGCCAGAGGACAAAGGGCTTAAGCTTTTCCA 1816
QY 121 AGCTACAAACAAACGTCGATTTGGGGATTTGATGATGAGACGAGCTGCTCTCTG 180
|||
DB 1817 ATCAACAGACGAC--TTGATGCTTTGGGATTTTGCATGAGACGAGCTGCTCTCTG 1873
QY 181 ACTGCGGGAGAGCGGAGCAAAATTCGCGGAGATCAGCCGAAAGAGCAAGTGGTGAAT 240
|||
DB 1874 ACTGCAAGGAGAGCGGAGCAAAATTCGCGGAGATCAGCAAGCAAGTGGTGAAT 1933
QY 241 ATGCTGGGAGACGTGGAGAAATTCAAAGACAGCAAAACCTCATGATGAGCGTCAAG 300
|||
DB 1934 ATGCTGGGAGATGGGAGACATATAGCAAGTACGCAAACTCATGATGAGCGTCAAG 1993
QY 301 GGAATCCCATGACATCCGCGGAGCCGATGTGTGATGCTCTCTGAACTTGGAGAAATG 360
|||
DB 1994 GGAATCCCATGACATCCGCGGAGCCGATGTGTGATGCTCTCTGAACTTGGAGAAATC 2053
QY 361 AAGTTGAAACCCCGGAGATACAGATCATGAAAGAGAGGCAAGGTCATCTGAG 420
|||
DB 2054 AAGTTGAAACCCCGGAGATACAGATCATGAAAGAGAGGCAAGGTCATCTGAA 2113
QY 421 CACATCAAGCGCATGACCGGAGCGTACCGGAGCACTTAAGAGATATTTCTTCAAG 480
|||
DB 2114 CACATCAAGCAATCAACCTGACGTGAGAGAGCACTCCGGAACATGCTCTTCTTAAAG 2173
QY 481 GATCGTAAGGAACCAAGAGCGGGAACCTACTCAATCTCTCTGATATGAGAGTAT 540
|||
DB 2174 GATCGTAATGAGACCAAGAGAGGGAACCTATCTCTCTGATATGAGAGTAT 2233
QY 541 AACCCGAGAGTGGCTTACTGCAAGGACCTGAGGCAATGCGCGCTTGTCTCTCTAT 600
|||
DB 2234 AACCCGAGAGTGGCTTACTGCAAGGACCTGAGGCAATGCGCGCTTGTCTCTCTAT 2293
QY 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTCACTGTGTGGCCAGTGAAGGCACTCC 660
|||
DB 2294 CTGCTGAGAGAGAGCACTTCTGGGCACTGTGTCACTGTGTGGCCAGTGAAGGCACTCC 2353
QY 661 CTGCAAGGATTTTCAACAGCCCAATGCGGGGACCGTCCAGGGGCTCCAAAGCAACAGGAG 720
|||
DB 2354 CTGCGAGATTTTCAACAGCCCAATGCGGGGACCGTCCAGGGGCTCCAAAGCAACAGGAG 2413
QY 721 CATGTGGTACGCAAGTCAACCAAGACATGAGGGGATGAGGACAAAGAAATCTATGT 780
|||
DB 2414 CATGTGGTACGCAAGTCAACCAAGACATGAGGGGATGAGGACAAAGAAATCTATGT 2473
QY 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTGACGGGATCTCTCGGG 840
|||
DB 2474 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTGACGGGATCTCTCGGG 2533
QY 841 CTCAACCTGCGCTGTGGGACGTGTATCTGCTAGAAAGGCAACAGCGTTGATGCGGATA 900
|||
DB 2534 CTCAACCTGCGCTGTGGGACGTGTATTTGGTGAAGAGAGAGGCTGTGATGCGGATA 2593
QY 901 ACAAGAAATGCGCTTTAAGTTTCAAGCAAGAGGCGCTTCAAGAAAGCTCAAGTGTGCGCG 960

Db	2554	ACGACGATGCTCTTAAGGTTGAGGAAAGCGCCTCATTGAAGACATTCAGGCTGGCCGTG	2653
Oy	961	TGGGCACTTTTGTGCACCGGTTCTGTGANTACTTGGGCGAGGAGTAGAGACACTGTGCTC	1020
Db	2654	TGGGCACTGTGGCGGACCAATTCTTCGATACCTGGGCGCATGAACGATGACACCGTGTCTC	2713
Oy	1021	AAGCATCTTAGGGCTCTATGAGAACTAACAGAAAGCGAGGGGACCTTGACACCCCA	1080
Db	2714	AAGCATCTTAGGGCTCTACAGAAAGAAATTAACAAGGAAGCAAGGGGACCTTGACACCCCA	2773
Oy	1081	GCCAAACCCGAGCAAGGGTGTGGGATCCAGGCGCTGTGACGGGCTTCACTGTGGGGGAAG	1140
Db	2774	GCCAAACCCGAGCAAGGGTCTTGGGCACTCAGGCGCTGTGCGGCTTCACTGTGGGGGAAG	2833
Oy	1141	ACCCTCTGCAAGGGGGAAGGACGAGGCGCCCTTCAAGGCCACAGACCCGGTTCCGCGGCCC	1200
Db	2834	ACCCTCTGCAAGGGGATATAGGCAAGGCGCCCTCAGAGGCCACAGGCCAGTTCCAGCGGCC	2893
Oy	1201	ATTGGTAGGCTCCCGGCGCAAGGGCACTGGTCTTCAACCTGTCTGTGGGGGCT	1260
Db	2894	ATTGCTAGGCTTCCCGGCGCAAGGGCACTGTGTTTCAAGCGCTGTCTGTGGGGGCT	2953
Oy	1261	GTCCGGGAAGACACTTACCTGTGTGGGCACTCAGGGTGTGCCAGGCCGACCTTGACTAG	1320
Db	2954	GTCCGGGAAGACACTTACCTGTGTGGGCACTCAGGGTGTGCCAGGGCTTGACTAG	3013
Oy	1321	GAGGACCTCAGGGTCTCTGAGATCTCTGACATGGAACCTCAGTCCCGCTCCCAAG	1380
Db	3014	GAGGACCTCAGGGTCTCTGAGATCTCTGAGATGGAAGTAAATGCCCGGCTCCCAAG	3073
Oy	1381	GACCTGACGTAGAAGGCGCTTGTGTTCCGCAATATGATTTCAAGCAGAGCTGTGGGTC	1440
Db	3074	GACCTGACGTATAGGGGGGCTTGTGTTCCCAATATGATTTTGAACGAGCTGTGGGTC	3133
Oy	1441	CGTGCATATCCAGAGAGACACAGCTGAGCCCTCTGTGGCAGGCTGAACACCTTGGGAG	1500
Db	3134	CGTGCATATCCAGAGAGACACAGCTGAGCCCTCTGTGGCAGGCTGAACCTTGGGAG	3193

```

RESULT 14
US-10-450-763-24016
; Sequence 24016, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 24016
LENGTH: 8180
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (4369)..(6126)
OTHER INFORMATION: 97% homologous to Homo sapiens oncogene, accession number
US-10-450-763-24016
OTHER INFORMATION: X63546, Smith-Waterman Score=3063.

```

Query Match	50.9%	Score 999;	DB 9;	Length 8180;
Best Local Similarity	77.8%;	Pred. No. 3.3e-305;		
Matches 1332;	Conservative	0;	Mismatches 165;	Indels 214;
				Gaps 2;

QY	1	ATGACGCTGTGTAAGAGTTCGCGGACATTGTGTGGGCAACAAGCCGAGAGACATCATTAATG	60
Db	1697	ATGGACATGTGTAGAGAAATGTCAAGATTGTGTGACAGGACACGAGGCGGAAAGACATCTTAATG	1756
QY	61	AAATTCGAAAGGGGCAACCGAGCTGTGGGCTGGCAGAGGCAAGAGGGGCTTAAGCCTTTTGG	120
Db	1757	AAGTATGACAAAGGACACCGAGCTGTGGGCTGGCAGAGGCAAGAGGGGCTTGAGCCCGTTTGA	1814
QY	121	AGCTTCAACAACAAAGTGCATCATTTTGGGGATTGTACATGTAGACGAGACTGCTCTCTG	180
Db	1817	A---TCAACAGACGATGATTGCTGTTTGGCATTTTGCATTTGATGAGACGAGACTGCTCTCTG	1873
QY	181	ACTGCGCGGAGGCGGAGCAAAATTCGCGCGGAGATCAGCCGAAAGACAGTGGGTGAT	240
Db	1874	ACTGCAACGGGAGGCGGAAAAATTCGCGCGGAGATGACACGAAACGACAACTGTGATGAA	1933
QY	241	ATGCTGGGAGACTGGGAGAAATATCAAAAGCAGCAAGAACTCATGTATCGAGCCGTACAG	300
Db	1934	ATGCTGGGAGATGGGAGACATATATAGACAGTGTCAAACTCATGTATCGAGTGTACAG	1993
QY	301	GGAATGCCCATGAACATCCGCGGCGCCGATGTGTCAAGTCTCTGAACATTTAGAGAAATG	360
Db	1994	GGAATTTCCATGAACATCCGCGGCGCCGATGTGTCAAGTCTCTCTGAACTTCAAGAAATC	2055
QY	351	AAGTTGAAAAACCCCGAAGATACCAATCATGTAAAGAGAGAGGCGAAGGTATCTGAG	420
Db	2054	AAGTTGAAAAACCCCGAAGATATCAAGATCATGAAGAGAGAGGCGCAAGAGGTATCTGAA	2113
QY	421	CACATCCAGCCCATTCGACCCGGAGCGTAAAGCCGGACATTAAAGAAACATATTTCTTCAAG	480
Db	2114	CACATCCACCACTTCGACCTGAGCGTGAAGCGACTCTCCGAAACATGTCTTTTAAGG	2173
QY	481	GATGATATCGGAAACCAAGCAGCGGAACTATCCACATCTCTCTGGCATATAGAGATAT	540
Db	2174	GATGATATGAGCCCAAGCAGAGGAACTATTTACATCTCTCTGGCATATAGAGATAT	2233
QY	541	AACCCGAGGTGGGCTACTGAGGAGGACCTGAGCCACATGCGCGCTTGTTCCTCTCTAT	600
Db	2234	AACCCGAGGTGGGCTACTGAGGAGGACCTGAGCCACATCACCGCTTGTTCCTCTTAT	2292
QY	601	CTTCTTGAGGAGATGCATTTCTGGGCACTGGTGCAGCTGCTGGCAGTGAAGAGCATCTC	660
Db	2294	CTGCTTGAGGAGAGCGCATTTCTGGGCACTGGTGCAGCTGCTGGCAGTGAAGAGCATCTC	2353
QY	661	CTGCAGGGATTTCAAGCCCAAAATGGGCGGAGCCGTCGAGGGGCTTCMAAGCAACAGAG	720
Db	2354	CTGCAGGATTTCAAGCCCAAAATGGTGGGCACTGTCAGGGGCTTCMAAGCAACAGAGAG	2413
QY	721	CATGTGTAGCCAGTCACAACCCCAAGCACTAGGGGCACTCAGCAACAAGAAAGATCTATGT	780
Db	2414	CATGTGTATCCCAAGTCACAACCCCAAGCACTATGGGATCTCAGCAACAAGAAAGTCTATGC	2473
QY	781	GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATTGACGGGATTTCTCTGGG	840
Db	2474	GGGCAGTGTGCTGTGTTAGGCTGCTCTCTCCGAACTCATGTGACGGGATTTCTCTGGG	2533
QY	841	CTCAACCCGCGCTGTGGGAGCTGTATCTGGTAAAGGCGCAAGAGGGTGTGATGCCATTA	900
Db	2534	CTCAACCTGCGCTGTGGAGCTGTATTTGGTGAAGGAGAAACAGTGTGTATGATCAATTA	2593
QY	901	ACAAGAAATCGCTTTAAGTTCAAGCAGAGAGCGCTCTCAAGAAAGTCCAGGTGTGGCCG	960
Db	2594	ACCAAGCAATGTCTTTAAGTTCAAGCAGAGAGCGCTCTCAAGAAAGTCCAGGTGTGGCCG	2653
QY	961	TGGGCAAGTTTTCGAAACCGGTTTGTGTATCTCTGGGCAAGGAGTGAAGACATGTGTCTC	1020
Db	2654	TGGGCAAGTCTTCGCGAAACCAATTTGTGATPACTCTGGGCAATGAACGATGACACCTGTCTC	2713
QY	1021	AAGCATCTTGAAGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGAGCTGCAACCCC--	1078
Db	2714	AAGCATCTTGAAGGCTCTTACGAAAGAACTAAACAAGAAAGCAGGGGAGCTGCAACCCC	2773
QY	1079	-----	1078

Db 2774 GGCCCAACAGCCCTGGGACGAAGGTGTGTGACAGGAAGCCCCAGCAGTCTGAACCTTG 2833
Qy 1079 ----- 1078
Db 2834 GGGGAGTCCAGAGACCAACCACAGTCCCAAGGCTTCCCATGCGAGGACACAC 2893
Qy 1079 ----- 1078
Db 2894 ACCCTCCCTGTGGGATGACGACGCTACAGGCGTGTGTGATGTGACGACGAGGGCC 2953
Qy 1079 -----CAGCCAAACCCGAGCAGGGGTGTGTGCAATC 1109
Db 2954 ACACAGAGACCCCAAGGACTCCAGATGACAGCCAAAGCGAGAGAGGCTCTTGGCAC 3013
Qy 1110 CAGGCTGTGCGGCTTCAAGTGTGCGGGAAGACCTCTGCAAGGGGGGACAGGAGGCCCC 1169
Db 3014 CAGGCTGTGCGGCTTCAAGTGTGCGGGAAGACCTCTGCAAGGGGGATAGGAGGCCCC 3073
Qy 1170 TCCAGGCCCCACAGCCCGGTTCCGCGGCCCATTTGTGACCTTCCCGCACGCGGAC 1229
Db 3074 TCCAGGCCCCACAGCCCGGTTCCAGCGGCCCATTTGTGACCTTCCCGCATGGGATC 3133
Qy 1230 TCGTTCTTTCACACCTGTCTGTGTGGGCTGTCCGGGAGACACTTACCTGTGGGAC 1289
Db 3134 TCGTTTTCACAGCCCTGTCTGTGTGGGCTGTCCGGGAGACACTTACCTGTGGGAC 3193
Qy 1290 TCAGGGTGTGCGGCCCCGCGCTGTGCTCAGGAGAGACCTGAGGTTCTGGAATTCCT 1349
Db 3194 TCAGGGTGTGCGGCCCCGCGCTGTGCTCAGGAGAGACCTGAGGTTCTGGAATTCCT 3253
Qy 1350 GCACTGGAACCTCAGTCCCGCTCCCAACGGAACCTGACGTAGAGGCCCCGTTGTTCCG 1409
Db 3254 GAGTGGAGATGATGCTGCGGCTCCCAACGGAACCTGATATAGGGGCCCTTGTGTTCC 3313
Qy 1410 CCATTATGATTTCAAGACAGAGCTGTGCTGTGCTGCTATCCAGAGAGACAGCTGAC 1469
Db 3314 CCATTATGATTTTAAACGAGCTGTGCTGTGCTGCTATCCAGAGAGACAGCTGAC 3373
Qy 1470 CCCCTGCTGGAGGCTGAACACCTGCGGAG 1500
Db 3374 CACCTGCTGGAGGCTGAACACCTGCGGAG 3404

RESULT 15

US-10-283-975A-272
; Sequence 272, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283, 975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340, 938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338, 997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340, 081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341, 012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 8201
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-272

Query Match 50.9%; Score 999; DB 7; Length 8201;
Best Local Similarity 77.8%; Pred. No. 3.36-305;
Matches 1332; Conservative 0; Mismatches 165; Indels 214; Gaps 2;

Qy 1 ATGACCTGTGTAAGAGTCCGCGGACAGTTGGTGGCAACAAGCGAGAGACATCATATATG 60
Db 1697 ATGACATGTGTAGAGATGACAGATAGTTTCCAGGACAGGAGGGAGACATATCTATG 1756
Qy 61 AAATACGAAAAGGACACCGAGCTGGGCTCCAGAGAGCAAGGGGCTTAAGCTTTTGA 120
Db 1757 AAGTATGACAAAGGACACCGAGCTGGGCTCCAGAGAGCAAGGGGCTTGGAGCCGTTGA 1816
Qy 121 AGCTACAAACAACCGTCATCTTTGGGGATTGTACATGACGAGGCTGCTCTG 180
Db 1817 A--TCAACGACAGCATTTGTTTGGCATTTTGTGATGACGAGGCTGCTCTG 1873
Qy 181 ACTGCGGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAGCAAGTGGTGGAT 240
Db 1874 ACTGCAAGGAGCGGAGAAATTCGCGGAGATGACAGAACGAACTGGATGGA 1933
Qy 241 ATGCTGGAGACTGGGAAATACAAAAGCAGCAAGAAAGCTCATGATCGAGCTACAG 300
Db 1934 ATGCTGGAGAAATGGGAGACATATTAAGCAGTAGCAAACTCATAGTCAAGTACAG 1993
Qy 301 GGAATGCCATGAAACATCCGGGGCCGATGTGTCAAGTCTCTGAAATTTGAGAAATG 360
Db 1994 GGAATTCATGAAACATCCGGGCGCGTGTGTGAGTCTCTGAAATTTGAGAAATC 2053
Qy 361 AAGTTGAAAAACCCGGAAGATACAGATCATGAAAGGAGCAAGAGTATCTGAG 420
Db 2054 AAGTTGAAAAACCCGGAATATACAGATCATGAAAGGAGCAAGAGTATCTGAA 2113
Qy 421 CACATCCAGGCAATCCAGCCGGAAGTACAGATCATGAAAGGAGCAATTAAGAACTATTTCTGAG 480
Db 2114 CACATCCAGCAATCCAGCTGGAAGTACAGATCATCTCCGGAACCAATGTCTTTAGG 2173
Qy 481 GATGATAGGAAACCAAGCGGGAATCTACATCTCTGAGATATGAGAGAT 540
Db 2174 GATGATATGAGCAAGCAAGAGAACTATTTACATCTCTGAGCTATTTGAGATAT 2233
Qy 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCAATCCGCTGTTGCTCTCTAT 600
Db 2234 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCAATCCGCTGTTGCTCTCTAT 2293
Qy 601 CTTCCTGAGAGAGATGATTTCTGGGCACTGTGTGAGCTGTGGCAGTGAAGGACTCC 660
Db 2294 CTTCCTGAGAGAGATGATTTCTGGGCACTGTGTGAGCTGTGGCAGTGAAGGACTCC 2353
Qy 661 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTCAGGGGCTCAAGACCAACGAGAG 720
Db 2354 CTGCAAGGATTTCAAGCCCAATGGGAGCTGCAAGGAGCTCAAGACCAACGAGAG 2413
Qy 721 CATGTGTAGCCAGTCAACCAAGACCATGGGCAATCAGAGCAAGAAAGTCTATGT 780
Db 2414 CATGTGTAGCCAGTCAACCAAGACCATGGGCAATCAGAGCAAGAAAGTCTATG 2473
Qy 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTTGATGACGGGATCTTCTGGG 840
Db 2474 GGGCAGTGTCTCTGTAGGCTGCTCTCCGAAACCTGATTAACGGGATCTTCTGGG 2533
Qy 841 CTGACCTGCGCTGTGGGACGTGTATGTGTGAAAGGAGCAAGGCTGTGATGCGGATA 900
Db 2534 CTGACCTGCGCTGTGGGACGTGTATGTGTGAAAGGAGCAAGGCTGTGATGCGGATA 2593
Qy 901 ACAAGATCCGCTTTAAGTTTCAAGAAAGCGCTCAGAAAGCTCAGAGTGTGGCCG 960
Db 2594 ACAAGATTCGCTTTAAGTTTCAAGAAAGCGCTCAGAAAGCTCAGAGTGTGGCCG 2653
Qy 961 TGGGCAAGTTTTCGAACCGGTTCTGTATACCTGGGCAAGGATGAGCACTGTGCTC 1020
Db 2654 TGGGCAAGTCTGGGAAACCAATTTCTGATACCTGGGCAAGGATGAGCACTGTGCTC 2713
Qy 1021 AAGCATCTTAAGGCGCTTATGAAAGAACTTAAGAAAGAGGAGGAGCTGGCAACCC-- 1078
Db 2714 AAGCATCTTAAGGCGCTTATGAAAGAACTTAAGAAAGAGGAGGAGCTGGCAACCCCA 2773

QY 1079 ----- 1078
Db 2774 GGCCCAACAGCCCTGGAGAGAGTGTGTGGCAAGAACCCCAAGCTTGAACCTTG 2833
QY 1079 ----- 1078
Db 2834 GGGGAGTCCAGAGAGCACCAACATGCCCAAGGCTTCCCATGCCAGGAGACAC 2893
QY 1079 ----- 1078
Db 2894 ACCCTCCCTCTGGATCAGACAGTACAGGCGTGTCTCATGTTCAGACACAGGAGCC 2953
QY 1079 ----- CAGCCAAACCCAGCAAGGCTGTGGATC 1109
Db 2954 ACACAGAGACCCCAAGACTCCAGAGATGCAAGCCAAACGAGCAAGGCTCTTGGCACC 3013
QY 1110 CAGGCTGTGCCGCTTCAAGTGGCGGAAAGACCTCTGCAAGGGGACAGGAGGCC 1169
Db 3014 CAGGCTGTGCCGCTTCAAGTGGGAAAGACCTCTGCAAGGGGATAGGAGGCC 3073
QY 1170 TCAGAGCCCAACAGCCCGTTCGCGGCTCATTTGTCACTTCCCGCACGGGACC 1229
Db 3074 TCAGAGCCCAACAGCCCGTTCGCGGCTCATTTGTCTCAGCTTCCCGCATGGGCATC 3133
QY 1230 TCGTCTTCAACACCCGTCTGTGGGGGCTGTCGCGGAAAGACACCTACCTGTGGGCAC 1289
Db 3134 TCGTCTTCAACACCCGTCTGTGGGGGCTGTCGCGGAAAGACACCTACCTGTGGGCAC 3193
QY 1290 TCAGGCTGTGCCAGCCCGCTCTGCTCAGGAGAGACCTCAGGTTCTTGAGATTCT 1349
Db 3194 TCAGGCTGTGCCAGCCCGCTCTGCTCAGGAGAGACCTCAGGTTCTTGAGATTCT 3253
QY 1350 GCAGTGGAACCTCATGCCCGCTCCCAACGACCTGACGTAGAGGCGCTTGTTCG 1409
Db 3254 GAGTGGAAATCATGCCCGCTCCCAACGACCTGACGTATAGGGGCGCTTGTTCG 3313
QY 1410 CCATTATGATTTGACAGAGCTGTGGTCCGTGCATATCCAGAGAGACCAAGCTGCG 1469
Db 3314 CCATTATGATTTGAGAGGAGCTGTGGTCCGTGCATATCCAGAGAGACCAAGCTGCG 3373
QY 1470 CCCCTGCTGCAAGCTGAACACCCCTGCGAG 1500
Db 3374 CACTGTGCTGCAAGCTGAACACTGCGGAGAG 3404

Search completed: April 5, 2006, 14:28:39
Job time : 1629 secs

2006 APR 06 16:21:14

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 14:20:23 ; Search time 1321 Seconds

(without alignments)
5947.837 Million cell updates/sec

Title: US-10-071-838-1

Perfect score: 1964

Sequence: 1 acggacgctgtagaggtcgc.....aaatgtctctgtgaatg 1964

Scoring table: IDENTITY_NUC

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.New:*
1: /SIDS5/prodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/prodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/prodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/prodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/prodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/prodata/1/pubpna/US09_NEW_PUB.seq:*
7: /SIDS5/prodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /SIDS5/prodata/1/pubpna/US10_NEW_PUB.seq2:*
9: /SIDS5/prodata/1/pubpna/US10_NEW_PUB.seq3:*
10: /SIDS5/prodata/1/pubpna/US10_NEW_PUB.seq4:*
11: /SIDS5/prodata/1/pubpna/US11_NEW_PUB.seq2:*
12: /SIDS5/prodata/1/pubpna/US11_NEW_PUB.seq3:*
13: /SIDS5/prodata/1/pubpna/US11_NEW_PUB.seq4:*
14: /SIDS5/prodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/prodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	358.6	18.3	378	6	US-09-925-065A-506682, Sequence 506682,
C 2	358.6	18.3	378	6	US-09-925-065A-506683, Sequence 506683,
C 3	301.6	15.4	186442	14	US-11-121-086-104, Sequence 104, App
C 4	300.6	15.3	477	6	US-09-925-065A-198638, Sequence 198638,
C 5	300.2	15.3	477	6	US-09-925-065A-198636, Sequence 198636,
C 6	300.2	15.3	477	6	US-09-925-065A-198637, Sequence 198637,
C 7	300.2	15.3	477	10	US-10-301-480-287417, Sequence 287417,
C 8	300.2	15.3	477	10	US-10-301-480-900827, Sequence 900827,
C 9	297.2	15.1	468	6	US-09-925-065A-173389, Sequence 173389,
C 10	297.2	15.1	477	10	US-10-301-480-264964, Sequence 264964,
C 11	297.2	15.1	477	10	US-10-301-480-878373, Sequence 878373,
C 12	294.2	15.0	476	10	US-10-301-480-285677, Sequence 285677,
C 13	294.2	15.0	476	10	US-10-301-480-899086, Sequence 899086,
C 14	294	15.0	479	6	US-09-925-065A-201701, Sequence 201701,
C 15	293.8	15.0	469	6	US-09-925-065A-198635, Sequence 198635,
C 16	293.8	15.0	469	10	US-10-301-480-287417, Sequence 287417,
C 17	293.8	15.0	469	10	US-10-301-480-900826, Sequence 900826,
C 18	293.8	15.0	479	10	US-10-301-480-294544, Sequence 294544,

C 19	293.8	15.0	479	10	US-10-301-480-294545, Sequence 294545,
C 20	293.8	15.0	479	10	US-10-301-480-907953, Sequence 907953,
C 21	293.8	15.0	479	10	US-10-301-480-907954, Sequence 907954,
C 22	293.6	14.9	459	10	US-10-301-480-264965, Sequence 264965,
C 23	293.6	14.9	459	10	US-10-301-480-878374, Sequence 878374,
C 24	293	14.9	465	6	US-09-925-065A-173390, Sequence 173390,
C 25	292.6	14.9	475	6	US-09-925-065A-173388, Sequence 173388,
C 26	292.6	14.9	479	10	US-10-301-480-294543, Sequence 294543,
C 27	292.6	14.9	479	10	US-10-301-480-907952, Sequence 907952,
C 28	292.2	14.9	477	10	US-10-301-480-294547, Sequence 294547,
C 29	292.2	14.9	472	6	US-09-925-065A-198614, Sequence 198614,
C 30	289.2	14.7	479	10	US-10-301-480-290201, Sequence 290201,
C 31	287	14.6	479	10	US-10-301-480-280202, Sequence 280202,
C 32	287	14.6	479	10	US-10-301-480-903610, Sequence 903610,
C 33	287	14.6	479	10	US-10-301-480-903611, Sequence 903611,
C 34	287	14.6	479	10	US-10-301-480-903611, Sequence 903611,
C 35	285.8	14.6	475	6	US-09-925-065A-585252, Sequence 585252,
C 36	285.4	14.5	479	10	US-10-301-480-290200, Sequence 290200,
C 37	285.4	14.5	479	10	US-10-301-480-903609, Sequence 903609,
C 38	281.8	14.3	467	6	US-09-925-065A-206786, Sequence 206786,
C 39	281.8	14.3	467	6	US-09-925-065A-206787, Sequence 206787,
C 40	281	14.3	473	6	US-09-925-065A-201700, Sequence 201700,
C 41	281	14.3	473	6	US-09-925-065A-201702, Sequence 201702,
C 42	280.6	14.3	467	6	US-09-925-065A-206784, Sequence 206784,
C 43	280.6	14.3	467	6	US-09-925-065A-206785, Sequence 206785,
C 44	280.2	14.3	465	6	US-09-925-065A-118026, Sequence 118026,
C 45	280.2	14.3	465	6	US-09-925-065A-206789, Sequence 206789,

ALIGNMENTS

RESULT 1
US-09-925-065A-506682/c
Sequence 506682, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Meng David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 506682
LENGTH: 378
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-506682

Query Match 18.3%; Score 358.6; DB 6; Length 378;
Best Local Similarity 99.7%; Pred. No. 3.8e-89;
Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1079 CAGCAAAACCCAGAGAGGTCGTGCGATCAGAGCCCTTGCAGCTTCACTGCGGGA 1138
DB 359 CAGCAAAACCCAGAGAGGTCGTGCGATCAGAGCCCTTGCAGCTTCACTGCGGGA 300
QY 1139 AGACCTTGCAGAGGAGGAGCAGAGCAGGCCCCCTTGCAGAGCCACAGCCCGGTTCCGCGGC 1198
DB 299 AGACCTTGCAGAGGAGGAGCAGAGCAGGCCCCCTTGCAGAGCCACAGCCCGGTTCCGCGGC 240

```
Qy 1199 CCAATTGGTCACTTCCCGGACGAGGCACTCGTCTTCCACACCCTGTCTGGTGGG 1258
|||
Db 239 CCAATTGGTCACTTCCCGGACGAGGCACTCGTCTTCCACACCCTGTCTGGTGGG 180
Qy 1259 CTGTCCGGGAAGACCTTACCTTGTGGCACTCAGGGGTGTCACAGCCCGGCTTGCTC 1318
|||
Db 179 CTGTCCGGGAAGACCTTACCTTGTGGCACTCAGGGGTGTCACAGCCCGGCTTGCTC 120
Qy 1319 AGGAGGACCTCAGGGTTCCTGAGATTCTCGAGTGAATCCAGTCCCGCTCCCAA 1378
|||
Db 119 AGGAGGACCTCAGGGTTCCTGAGATTCTCGAGTGAATCCAGTCCCGCTCCCAA 60
Qy 1379 CGGACCTGAGAGTGAAGGCGCTTGTCCGCACTTATGATTTCAGACAGAGCTGTGG 1437
|||
Db 59 CGGACCTGAGAGTGAAGGCGCTTGTCCGCACTTATGATTTCAGACAGAGCTGTGG 1

RESULT 2
US-09-925-065A-506683/C
; Sequence 506683, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 506683
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-506683

Query Match 18.3%; Score 358.6; DB 6; Length 378;
Best Local Similarity 99.7%; Pred. No. 3.8e-89;
Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1079 CAGCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTCGAGCTTCACTGAGCGGA 1138
|||
Db 359 CAGCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTCGAGCTTCACTGAGCGGA 300
Qy 1139 AGACCTCTGCAAGGAGGAGGAGGAGGCTTCCAGGCCCAAGCCGAGTTCCGCGGC 1198
|||
Db 299 AGACCTCTGCAAGGAGGAGGAGGAGGCTTCCAGGCCCAAGCCGAGTTCCGCGGC 240
Qy 1199 CCAATTGGTCACTTCCCGGACGAGGCACTCTGTTCCACACCTGTCTGTGGTGGG 1258
|||
Db 239 CCAATTGGTCACTTCCCGGACGAGGCACTCTGTTCCACACCTGTCTGTGGTGGG 180
Qy 1259 CTGTCCGGGAAGACCTTACCTTGTGGCACTCAGGGGTGTCACAGCCCGGCTTGCTC 1318
|||
Db 179 CTGTCCGGGAAGACCTTACCTTGTGGCACTCAGGGGTGTCACAGCCCGGCTTGCTC 120
Qy 1319 AGGAGGACCTCAGGGTTCCTGAGATTCTCGAGTGAATCCAGTCCCGCTCCCAA 1378
|||
Db 119 AGGAGGACCTCAGGGTTCCTGAGATTCTCGAGTGAATCCAGTCCCGCTCCCAA 60
Qy 1379 CGGACCTGAGAGTGAAGGCGCTTGTCCGCACTTATGATTTCAGACAGAGCTGTGG 1437
|||
Db 59 CGGACCTGAGAGTGAAGGCGCTTGTCCGCACTTATGATTTCAGACAGAGCTGTGG 1
```

```
RESULT 3
US-11-121-086-104
; Sequence 104, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 104
; LENGTH: 186442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-104

Query Match 15.4%; Score 301.6; DB 14; Length 186442;
Best Local Similarity 80.0%; Pred. No. 3.2e-72;
Matches 389; Conservative 0; Mismatches 74; Indels 23; Gaps 2;

Qy 1502 GGGTGAAGATGGGCTTTCGTCGACCCGACGATGATTCGACAGGACCCCTTCAGAG 1561
|||
Db 90831 GAACTAGCTGCTCCCTTACAGTGCACACGACTGCTTCCAAACAGGCGACCCCTTCACAG 90890
Qy 1562 CTAGGAGCAACAGCCGTGTGCTCCCACTCAGGCGCTTCTGTGCGCTTCCTACTTGG 1621
|||
Db 90891 CTAGGAGCAACAGCCGTGTGCTCCCACTCAGGCGCTTCTGTGCGCTTCCTACTTGG 90950
Qy 1622 AAAGTTTCAGTTCCTTCGAGGCTTCTGAGAGCATCTGGGCGAGGCTCATGTGCTGATA 1681
|||
Db 90951 AAAGTTTCAGTTCCTTCGAGGCTTCTGAGAGCATCTGGGCGAGGCTCATGTGCTGATA 91010
Qy 1682 ATTTCCCTAGCTTAAACCAACCAAGGCTGCGCTTCGTTTATTTTGGTTAAAC 1741
|||
Db 91011 AGCACCAGAGTCCCAACCAACCAAGGCTGCGCTTCGTTTATTTTGGTTAAAC 91070
Qy 1742 TTATGAATGATTAAGA-----AAGATGACGCTTGAGAGA 1779
|||
Db 91071 TTATGAATGATTAAGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91130
Qy 1780 GATTCAAGATGGAACACACACGACCCCAATCAAAAGCAACATGCCAGCCCTCC 1839
|||
Db 91131 CATTCAAAAGATGAACACACGACCCCAATCAAAAGCAACATGCCAGCCCTCC 91190
Qy 1840 CAGACCCCGAGCCCGACGACCATGTTCTGAATTCGACGACCCGTGAGCTGACC-TT 1898
|||
Db 91191 CAGTCCCTGATGTCCTGCAACCAAGTGTCTGAATTCGACGACCTGAGCTGCTT 91250
Qy 1899 TGTACTTAACTCATGGAAGATTAACCTTCACTGTTTGAATTAATGTTTCTGTT 1958
|||
Db 91251 TGTACTTAACTCATGGAAGATTAACCACTTCTGTTTAAATTAATGTTTACTTT 91310
Qy 1959 GAAATG 1964
|||
Db 91311 GAAATG 91316

RESULT 4
US-09-925-065A-198638
; Sequence 198638, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827.135
```

;; CURRENT APPLICATION NUMBER: US/09/925,065A
;; CURRENT FILING DATE: 2001-08-08
;; PRIOR APPLICATION NUMBER: US 60/243,096
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 198638
;; LENGTH: 477
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-198638

Query Match
Best Local Similarity 15.3%; Score 300.6; DB 6; Length 477;
Best Local Similarity 98.7%; Pred. No. 5.3e-73;
Matches 303; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1658 TGGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGAGCTTCGCG 1717
DB 1 TGGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGAGCTTCGCGA 60
QY 1718 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGTGACGCTGAGA 1777
DB 61 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGTGACGCTGAGA 120
QY 1778 GAGATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGCCTCCAGCCCT 1837
DB 121 GAGATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGCCTCCAGCCCT 180
QY 1838 CCAGACACCCCGAGCCCGACGACCATGTTCTGAATTTCTAGCAGACCCGAGCCTGCT 1897
DB 181 CCAGACACCCCGAGCCCGACGACCATGTTCTGAATTTCTAGCAGACCCGAGCCTGCT 240
QY 1898 TTGTACTTTAACTGAGAGATTAACCTTCACTTCACTTCACTTCACTTCACTTCTGT 1957
DB 241 TTGTACTTTAACTGAGAGATTAACCTTCACTTCACTTCACTTCACTTCACTTCTGT 300
QY 1958 TGAATG 1964
DB 301 TGAATG 307

RESULT 5
US-09-925-065A-198636
;; Sequence 198636, Application US/09925065A
;; Publication No. US20040181048A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single
;; FILE REFERENCE: 108827.135
;; CURRENT APPLICATION NUMBER: US/09/925,065A
;; CURRENT FILING DATE: 2001-08-08
;; PRIOR APPLICATION NUMBER: US 60/243,096
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 198636

;; LENGTH: 477
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-198636

Query Match
Best Local Similarity 15.3%; Score 300.2; DB 6; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1658 TGGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGAGCTTCGCG 1717
DB 1 TGGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGAGCTTCGCGA 60
QY 1718 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGTGACGCTGAGA 1777
DB 61 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGTGACGCTGAGA 120
QY 1778 GAGATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGCCTCCAGCCCT 1837
DB 121 GAGATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGCCTCCAGCCCT 180
QY 1838 CCAGACACCCCGAGCCCGACGACCATGTTCTGAATTTCTAGCAGACCCGAGCCTGCT 1897
DB 181 CCAGACACCCCGAGCCCGACGACCATGTTCTGAATTTCTAGCAGACCCGAGCCTGCT 240
QY 1898 TTGTACTTTAACTGAGAGATTAACCTTCACTTCACTTCACTTCACTTCACTTCTGT 1957
DB 241 TTGTACTTTAACTGAGAGATTAACCTTCACTTCACTTCACTTCACTTCACTTCTGT 300
QY 1958 TGAATG 1964
DB 301 TGAATG 307

RESULT 6
US-09-925-065A-198637
;; Sequence 198637, Application US/09925065A
;; Publication No. US20040181048A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single
;; FILE REFERENCE: 108827.135
;; CURRENT APPLICATION NUMBER: US/09/925,065A
;; CURRENT FILING DATE: 2001-08-08
;; PRIOR APPLICATION NUMBER: US 60/243,096
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 198637
;; LENGTH: 477
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-198637

Query Match
Best Local Similarity 15.3%; Score 300.2; DB 6; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1658 TGGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGAGCTTCGCG 1717
DB 1 TGGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGAGCTTCGCGA 60
QY 1718 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGTGACGCTGAGA 1777
DB 61 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGTGACGCTGAGA 120

Db 61 TCCTGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGAGTCAGCTCGAGA 120
Qy 1778 GAGATTGAGATGGAACACGACCCGAGATCACAACCAACCAATGCCAGCCCT 1837
Db 121 GAGATTGAGATGGAACACGACCCGAGATCACAACCAACCAATGCCAGCCCT 180
Qy 1838 CCCAGCACCCCGACCCGACCAACCATCTGTGTAATTTGACGACACCGTGAAGCTGCT 1897
Db 181 CSCAGCACCCCGACCCGACCAACCATCTGTGTAATTTGACGACACCGTGAAGCTGCT 240
Qy 1898 TTGACTTTAACTGATGAGATGAGATTAACCTTCAAGTTTGAATTAATGTTTCTGCT 1957
Db 241 TTGACTTTAACTGATGAGATGAGATTAACCTTCAAGTTTGAATTAATGTTTCTGCT 300
Qy 1958 TGAATG 1964
Db 301 TGAATG 307

RESULT 7

US-10-301-480-287418/c
; Sequence 287418, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287418
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-287418

Query Match 15.3%; Score 300.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1658 TGGGCCAGGGCTCATGCTGATTAATTTCCCTAGGCTTAACAACCAAGCAAGCTTGGCG 1717
Db 477 TGGGCCAGGGCTCATGCTGATTAATTTCCCTAGGCTTAACAACCAAGCAAGCTTGGCG 418
Qy 1718 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGAGTCAGCTCGAGA 1777
Db 417 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGAGTCAGCTCGAGA 358
Qy 1778 GAGATTGAGATGGAACACGACCCGAGATCACAACCAACCAATGCCAGCCCT 1837
Db 357 GAGATTGAGATGGAACACGACCCGAGATCACAACCAACCAATGCCAGCCCT 298
Qy 1838 CCCAGCACCCCGACCCGACCAACCATCTGTGTAATTTGACGACACCGTGAAGCTGCT 1897
Db 297 CSCAGCACCCCGACCCGACCAACCATCTGTGTAATTTGACGACACCGTGAAGCTGCT 238
Qy 1898 TTGACTTTAACTGATGAGATGAGATTAACCTTCAAGTTTGAATTAATGTTTCTGCT 1957
Db 237 TTGACTTTAACTGATGAGATGAGATTAACCTTCAAGTTTGAATTAATGTTTCTGCT 178
Qy 1958 TGAATG 1964
Db 177 TGAATG 171

RESULT 8
US-10-301-480-900827/c

; Sequence 900827, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 900827
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-900827

Query Match 15.3%; Score 300.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1658 TGGGCCAGGGCTCATGCTGATTAATTTCCCTAGGCTTAACAACCAAGCAAGCTTGGCG 1717
Db 477 TGGGCCAGGGCTCATGCTGATTAATTTCCCTAGGCTTAACAACCAAGCAAGCTTGGCG 418
Qy 1718 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGAGTCAGCTCGAGA 1777
Db 417 TCCTCGTTTATTTTGGTTAACTTAAGAAATGTAATTAAGAAAGAGTCAGCTCGAGA 358
Qy 1778 GAGATTGAGATGGAACACGACCCGAGATCACAACCAACCAATGCCAGCCCT 1837
Db 357 GAGATTGAGATGGAACACGACCCGAGATCACAACCAACCAATGCCAGCCCT 298
Qy 1838 CCCAGCACCCCGACCCGACCAACCATCTGTGTAATTTGACGACACCGTGAAGCTGCT 1897
Db 297 CSCAGCACCCCGACCCGACCAACCATCTGTGTAATTTGACGACACCGTGAAGCTGCT 238
Qy 1898 TTGACTTTAACTGATGAGATGAGATTAACCTTCAAGTTTGAATTAATGTTTCTGCT 1957
Db 237 TTGACTTTAACTGATGAGATGAGATTAACCTTCAAGTTTGAATTAATGTTTCTGCT 178
Qy 1958 TGAATG 1964
Db 177 TGAATG 171

RESULT 9

US-09-925-065A-173389
; Sequence 173389, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 173389
LENGTH: 468
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-173389

Query Match 15.1%; Score 297.2; DB 6; Length 468;
Best Local Similarity 98.4%; Pred. No. 4.7e-72;
Matches 299; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1661 GCCAGGGCTCATGGCTGATTAATTCCTTAAGGCTTAACAACCCAGACAGCTTCGCGTCC 1720
DB 1 GCCAGGGCTCATGGCTGATTAATTCCTTAAGGCTTAACAACCCAGACAGCTTCGCGATCC 60
QY 1721 TCGTTTATTTTGGTTAACTTAATGATAAATGATTAAAGAAAGTGCAGCTTGAAGAG 1780
DB 61 TCGTTTATTTTGGTTAACTTAATGATAAATGATTAAAGAAAGTGCAGCTTGAAGAG 120
QY 1781 ATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGATGCCAGCCCTCCC 1840
DB 121 ATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGATGCCAGCCCTCCC 180
QY 1841 AGACCCCCAGCCCCAGACCATGTTCTGAATTTGACGACACCGTGAAGCTTGCCTTTG 1900
DB 181 AGACCCCCAGCCCCAGACCATGTTCTGAATTTGACGACACCGTGAAGCTTGCCTTTG 240
QY 1901 TACTTTAACTCATGGAAGATTAACCTTCACTTTTGAATTAATGTTTCTGTTGA 1960
DB 241 TACTTTAACTCATGGAAGATTAACCACTTCACTTTTGAATTAATGTTTCTGTTGA 300
QY 1961 AATG 1964
DB 301 AATG 304

RESULT 10

US-10-301-480-264964
Sequence 264964, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 264964
LENGTH: 477
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-264964

Query Match 15.1%; Score 297.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 4.7e-72;
Matches 299; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1661 GCCAGGGCTCATGGCTGATTAATTCCTTAAGGCTTAACAACCCAGACAGCTTCGCGTCC 1720
DB 1 GCCAGGGCTCATGGCTGATTAATTCCTTAAGGCTTAACAACCCAGACAGCTTCGCGATCC 60
QY 1721 TCGTTTATTTTGGTTAACTTAATGATAAATGATTAAAGAAAGTGCAGCTTGAAGAG 1780
DB 61 TCGTTTATTTTGGTTAACTTAATGATAAATGATTAAAGAAAGTGCAGCTTGAAGAG 120
QY 1781 ATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGATGCCAGCCCTCCC 1840
DB 121 ATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGATGCCAGCCCTCCC 180

QY 1841 AGACCCCCAGCCCCAGACCATGTTCTGAATTTGACGACACCGTGAAGCTTGCCTTTG 1900
DB 181 AGACCCCCAGCCCCAGACCATGTTCTGAATTTGACGACACCGTGAAGCTTGCCTTTG 240
QY 1901 TACTTTAACTCATGGAAGATTAACCTTCACTTTTGAATTAATGTTTCTGTTGA 1960
DB 241 TACTTTAACTCATGGAAGATTAACCACTTCACTTTTGAATTAATGTTTCTGTTGA 300
QY 1961 AATG 1964
DB 301 AATG 304

RESULT 11

US-10-301-480-878373
Sequence 878373, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 878373
LENGTH: 477
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-878373

Query Match 15.1%; Score 297.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 4.7e-72;
Matches 299; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1661 GCCAGGGCTCATGGCTGATTAATTCCTTAAGGCTTAACAACCCAGACAGCTTCGCGTCC 1720
DB 1 GCCAGGGCTCATGGCTGATTAATTCCTTAAGGCTTAACAACCCAGACAGCTTCGCGATCC 60
QY 1721 TCGTTTATTTTGGTTAACTTAATGATAAATGATTAAAGAAAGTGCAGCTTGAAGAG 1780
DB 61 TCGTTTATTTTGGTTAACTTAATGATAAATGATTAAAGAAAGTGCAGCTTGAAGAG 120
QY 1781 ATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGATGCCAGCCCTCCC 1840
DB 121 ATTCAGAGATGGAACACACAGACCCAGATCAACAAGCCAGATGCCAGCCCTCCC 180
QY 1841 AGACCCCCAGCCCCAGACCATGTTCTGAATTTGACGACACCGTGAAGCTTGCCTTTG 1900
DB 181 AGACCCCCAGCCCCAGACCATGTTCTGAATTTGACGACACCGTGAAGCTTGCCTTTG 240
QY 1901 TACTTTAACTCATGGAAGATTAACCTTCACTTTTGAATTAATGTTTCTGTTGA 1960
DB 241 TACTTTAACTCATGGAAGATTAACCACTTCACTTTTGAATTAATGTTTCTGTTGA 300
QY 1961 AATG 1964
DB 301 AATG 304

RESULT 12

US-10-301-480-285677/C
Sequence 285677, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

;; TITLE OF INVENTION: in the Human Genome
;; FILE REFERENCE: 108827.137
;; CURRENT APPLICATION NUMBER: US/10/301,480
;; PRIOR FILING DATE: 2002-11-21
;; PRIOR APPLICATION NUMBER: US 10/215,598
;; PRIOR FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: US 60/311,695
;; PRIOR FILING DATE: 2001-08-10
;; NUMBER OF SEQ ID NOS: 1226818
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 285677
;; LENGTH: 476
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-10-301-480-285677

Query Match 15.0%; Score 294.2; DB 10; Length 476;
Best Local Similarity 98.3%; Pred. No. 3.2e-71;
Matches 296; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1664 AGGCTCATGCTGGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 1723
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 476 AGGCTCATGCTGGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 417
QY 1724 TTTTATTTTGGTTAACTTATGAATAATGTATTAAGAAAGTGCAGCTCGAGAGATT 1783
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 416 TTTTATTTTGGTTAACTTATGAATAATGTATTAAGAAAGTGCAGCTCGAGAGATT 357
QY 1784 CAGAGATGAACACACCAAGACCCAGATCAACAAAGCCACATGCCCCCTCCAGC 1843
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 356 CAGAGATGAACACACCAAGACCCAGATCAACAAAGCCACATGCCCCCTCCAGC 297
QY 1844 ACCCCCAAGCCCAAGACCATGTTCTGAATTTCTGACGACCGTGAAGCTCTTTGTAC 1903
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 296 ACCCCCAAGCCCAAGACCATGTTCTGAATTTCTGACGACCGTGAAGCTCTTTGTAC 237
QY 1904 TTTAACTCATGGAAGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 1963
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 236 TTTAACTCATGGAAGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 177
QY 1964 G 1964
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 176 G 176

RESULT 13

US-10-301-480-899086/c
;; Sequence 899086, Application US/10301480
;; Publication No. US20060057564A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
;; FILE REFERENCE: 108827.137
;; CURRENT APPLICATION NUMBER: US/10/301,480
;; PRIOR FILING DATE: 2002-11-21
;; PRIOR APPLICATION NUMBER: US 10/215,598
;; PRIOR FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: US 60/311,695
;; PRIOR FILING DATE: 2001-08-10
;; NUMBER OF SEQ ID NOS: 1226818
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 899086
;; LENGTH: 476
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-10-301-480-899086

Query Match 15.0%; Score 294.2; DB 10; Length 476;
Best Local Similarity 98.3%; Pred. No. 3.2e-71;
Matches 296; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1664 AGGCTCATGCTGGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 1723

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 476 AGGCTCATGCTGGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 417
QY 1724 TTTTATTTTGGTTAACTTATGAATAATGTATTAAGAAAGTGCAGCTCGAGAGATT 1783
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 416 TTTTATTTTGGTTAACTTATGAATAATGTATTAAGAAAGTGCAGCTCGAGAGATT 357
QY 1784 CAGAGATGAACACACCAAGACCCAGATCAACAAAGCCACATGCCCCCTCCAGC 1843
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 356 CAGAGATGAACACACCAAGACCCAGATCAACAAAGCCACATGCCCCCTCCAGC 297
QY 1844 ACCCCCAAGCCCAAGACCATGTTCTGAATTTCTGACGACCGTGAAGCTCTTTGTAC 1903
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 296 ACCCCCAAGCCCAAGACCATGTTCTGAATTTCTGACGACCGTGAAGCTCTTTGTAC 237
QY 1904 TTTAACTCATGGAAGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 1963
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 236 TTTAACTCATGGAAGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 177
QY 1964 G 1964
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 176 G 176

RESULT 14

US-09-925-065A-201701/c
;; Sequence 201701, Application US/09925065A
;; Publication No. US20040181048A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single
;; FILE REFERENCE: 108827.135
;; CURRENT APPLICATION NUMBER: US/09/925,065A
;; PRIOR FILING DATE: 2001-08-08
;; PRIOR APPLICATION NUMBER: US 60/243,096
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 201701
;; LENGTH: 479
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-201701

Query Match 15.0%; Score 294; DB 6; Length 479;
Best Local Similarity 97.7%; Pred. No. 3.6e-71;
Matches 297; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1661 GCCAGGCTCATGCTGGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 1720
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 479 GCCAGGCTCATGCTGGATTAATTTCCCTAGGCTTAACCAACCAAGCTTGGCTCTCG 420
QY 1721 TCGTTTATTTTGGTTAACTTATGAATAATGTATTAAGAAAGTGCAGCTCGAGAGATT 1780
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 419 TCGTTTATTTTGGTTAACTTATGAATAATGTATTAAGAAAGTGCAGCTCGAGAGATT 360
QY 1781 ATTCAAGATGAACACACCAAGACCCAGATCAACAAAGCCACATGCCCCCTCCAGC 1840
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 359 ATTCAAGATGAACACACCAAGACCCAGATCAACAAAGCCACATGCCCCCTCCAGC 300
QY 1841 AGACCCCAAGCCCAAGACCATGTTCTGAATTTCTGACGACCGTGAAGCTCTTTGTAC 1900
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 299 AGACCCCAAGCCCAAGACCATGTTCTGAATTTCTGACGACCGTGAAGCTCTTTGTAC 240

Qy 1901 TACTTAACTGATGAGATTAATCTTACGTTTGAATAATGTTTCTGTGA 1960
|||
Db 239 TACTTAACTGATGAGATTAATCTTACGTTTGAATAATGTTTCTGTGA 180
Qy 1961 AATG 1964
|||
Db 179 AATG 176

RESULT 15

US-09-925-065A-198635
; Sequence 198635, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198635
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-198635

Query Match

Best Local Similarity 15.0%; Score 293.8; DB 6; Length 469;
Matches 295; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1666 GGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGCAGCTTGGCGTCTGTT 1725
|||
Db 1 GGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAGCAGCTTGGCGTCTGTT 60
Qy 1726 TTAATTTGGTTAACTTAATGAAATGTAATTAAGAAAGTGCAGCTCGAGAGATTCA 1785
|||
Db 61 TTAATTTGGTTAACTTAATGAAATGTAATTAAGAAAGTGCAGCTCGAGAGATTCA 120
Qy 1786 GAGATGGAACACACGACGACCCGAGATCACAAGCCACATGCCAGCCCTCCAGCAC 1845
|||
Db 121 GAGATGGAACACACGACGACCCGAGATCACAAGCCACATGCCAGCCCTCCAGCAC 180
Qy 1846 CCCGAGCCCGACGACATGTTCTGAATCTGACGACACCGGAGCTGCTTGTACTT 1905
|||
Db 181 CCCGAGCCCGACGACATGTTCTGAATCTGACGACACCGGAGCTGCTTGTACTT 240
Qy 1906 TAAACTCATGGAAGATTAATCACTTCAAGTTTGAATTAATGTTTCTGTTGAATG 1964
|||
Db 241 CAAACTCATGGAAGATTAATCACTTCAAGTTTGAATTAATGTTTCTGTTGAATG 299

Search completed: April 5, 2006, 14:42:32
Job time : 1321 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:18:49 ; Search time 811 Seconds
(without alignments)
16139.897 Million cell updates/sec

Title: US-10-071-838-1

Perfect score: 1964
Sequence: 1 atgagcgcgtgtagagtcgc.....aaatgttcctcgttgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.21.*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn20018:*
5: geneeqn20028:*
6: geneeqn20038:*
7: geneeqn20048:*
8: geneeqn20058:*
9: geneeqn20068:*
10: geneeqn20078:*
11: geneeqn20088:*
12: geneeqn20098:*
13: geneeqn20108:*
14: geneeqn20118:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1964	100.0	1964	6	ABSS4706
2	1964	100.0	1964	10	ADCC37382
3	1962.4	99.9	1964	10	ADCC37384
4	1948	99.2	2072	5	ABX71403
5	1944	99.0	2064	13	ADQ84596
6	1944	99.0	2064	13	ADQ86595
7	1929.8	98.3	1993	3	AAA15001
8	1929.8	98.3	1993	13	ADT88069
9	1902	96.8	2304	10	ADCC37380
10	1900.4	96.8	2304	10	ADCC37380
11	1836.4	93.5	2084	6	ABZ11810
12	1836.4	93.5	2084	12	ADMA44328
13	1812.4	92.3	2072	6	ABZ11809
14	1812.4	92.3	2072	12	ADMA44327
15	1809	92.1	2647	10	ADCC37388
16	1807.4	92.0	2647	10	ADCC37386
17	1807.4	92.0	2647	11	ADMO1991
18	1794	91.3	2146	6	ABSS4707
19	1750	89.1	1862	6	ABSS4708

20	1638.8	83.4	1752	6	ABSS78751	ABSS78751 DNA encod
21	1220	62.1	7856	5	AA885531	AA885531 DNA encod
22	1220	62.1	7856	5	AA888223	AA888223 DNA encod
23	1220	62.1	7878	6	AB164684	AB164684 Stomach c
24	1218.4	62.0	2881	13	ACN42882	ACN42882 Human dia
25	1128.4	57.5	3799	13	ACN42884	ACN42884 Human dia
26	1127.6	57.4	3901	13	ACN42883	ACN42883 Human dia
27	999	50.9	8180	5	AA888212	AA888212 DNA encod
28	999	50.9	8201	2	AA112170	AA112170 pUG4-5-CD
29	999	50.9	8201	10	AD885053	AD885053 Farnesyl
30	999	50.9	8284	5	AA885535	AA885535 DNA encod
31	999	50.9	8408	4	AA157848	AA157848 Human pol
32	999	50.9	8420	4	AA157847	AA157847 Human pol
33	999	50.9	8435	4	AA159632	AA159632 Human pol
34	999	50.9	8435	4	AA159634	AA159634 Human pol
35	999	50.9	8435	10	ADCC31883	ADCC31883 Human nov
36	882.8	44.9	9804	4	AA105173	AA105173 Human car
37	882.8	44.9	9804	4	AA105171	AA105171 Human rep
38	882.8	44.9	9804	4	AB198058	AB198058 Human tes
39	882.8	44.9	9804	10	AD847023	AD847023 Human car
40	882.8	44.9	9804	13	ADJ08441	ADJ08441 Human car
41	879.6	44.8	9404	4	AA536326	AA536326 Human car
42	879.6	44.8	9404	4	AA105171	AA105171 Human rep
43	879.6	44.8	9404	4	AB198058	AB198058 Human tes
44	879.6	44.8	9404	10	AD847020	AD847020 Human car
45	879.6	44.8	9404	13	ADJ08438	ADJ08438 Human car

ALIGNMENTS

RESULT 1	ABSS4706	ABSS4706 standard; CDNA; 1964 BP.
ID	ABSS4706	standard; CDNA; 1964 BP.
XX	ABSS4706;	
AC	ABSS4706;	
XX	28-NOV-2002 (first entry)	
DT	28-NOV-2002 (first entry)	
XX	CDNA encoding human PRC17 protein.	
DE	CDNA encoding human PRC17 protein.	
XX	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;	
KW	breast cancer; cytostatic; chromosome 17q11-12; gene; ss.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	CDS	1..1650
FT		/*tag= a
FT		/product= "Human PRC17 protein"
XX	WO200262958-A2.	
PN	WO200262958-A2.	
XX	15-AUG-2002.	
PD	15-AUG-2002.	
XX	08-FEB-2002; 2002WO-US003457.	
PF	08-FEB-2002; 2002WO-US003457.	
XX	08-FEB-2001; 2001US-0267615P.	
PR	08-FEB-2001; 2001US-0267615P.	
XX	(TUL-) TULARIK INC.	
PA	(TUL-) TULARIK INC.	
XX	Li J, Powers S, Xiang P, Peng Y;	
PI	Li J, Powers S, Xiang P, Peng Y;	
XX	WPI; 2002-706902/76.	
DR	P-PSDB; ABG70736.	
XX	P-PSDB; ABG70736.	
PT	Novel isolated PRC17 polypeptide useful- <u>diagnostically</u> or <u>prognostically</u>	
PT	to detect diseases or conditions associated with altered PRC17 activity	
PT	or expression relative-to-normal, for example-cancer.	
XX	Claim 24; Page 62; 78pp; English.	
PS	Claim 24; Page 62; 78pp; English.	
XX	The present invention relates to a new PRC17 polypeptide. The invention	

CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present nucleic acid sequence represents the human PRC17
CC located on chromosome 17q11-12. This sequence encodes the human PRC17
CC protein of the invention

XX Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 1964; DB 6; Length 1964;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGCGTGTAGAGTTCGCGGCGAGTTGTGGCACAGAGCGAGAGCATCATTTATG 60
    |||
DB 1 ATGAGCGTGTAGAGTTCGCGGCGAGTTGTGGCACAGAGCGAGAGCATCATTTATG 60
QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGGCAAGGGGCTTAAGCCTTTTCA 120
    |||
DB 61 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGGCAAGGGGCTTAAGCCTTTTCA 120
QY 121 AGCTACAAACAACAGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTG 180
    |||
DB 121 AGCTACAAACAACAGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTG 180
QY 181 ACTGCGCGGAGGCGAAGCAATTCGCGGAGATCAAGCCGAAGACAGTGGTGGAT 240
    |||
DB 181 ACTGCGCGGAGGCGAAGCAATTCGCGGAGATCAAGCCGAAGACAGTGGTGGAT 240
QY 241 ATGCTGGGAGACTGGGGAATTAACAAGACGAGAAAGCTCATAGATTCAGCGCTAAG 300
    |||
DB 241 ATGCTGGGAGACTGGGGAATTAACAAGACGAGAAAGCTCATAGATTCAGCGCTAAG 300
QY 301 GGAATGCGCATGACATCCGGGCGCCGATGTGTCACTCTCTGAAATTGAGGAATG 360
    |||
DB 301 GGAATGCGCATGACATCCGGGCGCCGATGTGTCACTCTCTGAAATTGAGGAATG 360
QY 361 AAGTTGAAAAAAGCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCACTGAG 420
    |||
DB 361 AAGTTGAAAAAAGCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCACTGAG 420
QY 421 CACATCCAGGCGATCGAACCGGGAAGTAAAGCGGGAATTAAAGAGCATATATTTTCA 480
    |||
DB 421 CACATCCAGGCGATCGAACCGGGAAGTAAAGCGGGAATTAAAGAGCATATATTTTCA 480
QY 481 GATGATACGGAACCAAGACGCGGGAATCTACATCTCTGATCTGATAGAGATAT 540
    |||
DB 481 GATGATACGGAACCAAGACGCGGGAATCTACATCTCTGATCTGATAGAGATAT 540
QY 541 AACCAGGAGTGGGCTACTGAGGAGACTGAGCAATCGCGCTTTGTTCTCTCTAT 600
    |||
DB 541 AACCAGGAGTGGGCTACTGAGGAGACTGAGCAATCGCGCTTTGTTCTCTCTAT 600
QY 601 CTTCTGAGAGAGATGCAATTTGGGCACTGTGTGAGCTGTGGCAAGTGAAGGCACTCC 660
    |||
DB 601 CTTCTGAGAGAGATGCAATTTGGGCACTGTGTGAGCTGTGGCAAGTGAAGGCACTCC 660
QY 661 CTGAGAGGATTTTCAAGCCCAATGCGGAGACGTTCCAGGGGCTTCAAGACCAAGAG 720
    |||
DB 661 CTGAGAGGATTTTCAAGCCCAATGCGGAGACGTTCCAGGGGCTTCAAGACCAAGAG 720
QY 721 CATGTGTAGCGCAGCTCAACCCCAAGACATGAGGAGATCAAGACAAAGATCTATGT 780
    |||
DB 721 CATGTGTAGCGCAGCTCAACCCCAAGACATGAGGAGATCAAGACAAAGATCTATGT 780
QY 781 GGGGAGGTTCCTCCGTTAGGCTGCTCATCCGAGATTTGATTTGACGGGATCTCTCGGG 840
    |||
DB 781 GGGGAGGTTCCTCCGTTAGGCTGCTCATCCGAGATTTGATTTGACGGGATCTCTCGGG 840
```

```
QY 841 CTCACCTGCGCTGTGGAGCTGTATCTGTGTGAAGGCAAGCGGCTGATGCCGATA 900
    |||
DB 841 CTCACCTGCGCTGTGGAGCTGTATCTGTGTGAAGGCAAGCGGCTGATGCCGATA 900
QY 901 ACAAGAAATGCGCTTTAAAGTTTCAGCAAGAGCGCTCAAGAGACGTCCAGGTGGCCG 960
    |||
DB 901 ACAAGAAATGCGCTTTAAAGTTTCAGCAAGAGCGCTCAAGAGACGTCCAGGTGGCCG 960
QY 961 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCGCAGGATGAGGACATGTGCTC 1020
    |||
DB 961 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCGCAGGATGAGGACATGTGCTC 1020
QY 1021 AAGCATTTAGGCGCTCTATGAAGAACTTAACAAGAGAGGAGGAGCTGCCACCCCA 1080
    |||
DB 1021 AAGCATTTAGGCGCTCTATGAAGAACTTAACAAGAGAGGAGGAGCTGCCACCCCA 1080
QY 1081 GCCAAACCCGAGCAAGGTTGTGGGATCCAGGCGCTGTGCGGCTTCAAGTGGCGGAG 1140
    |||
DB 1081 GCCAAACCCGAGCAAGGTTGTGGGATCCAGGCGCTGTGCGGCTTCAAGTGGCGGAG 1140
QY 1141 ACCCTTGCAGAGGGGACAGGCAAGGCCCTCCAGGCCCAACAGCCGAGTTCCGCGGCC 1200
    |||
DB 1141 ACCCTTGCAGAGGGGACAGGCAAGGCCCTCCAGGCCCAACAGCCGAGTTCCGCGGCC 1200
QY 1201 ATTTGTCAGCTTCCCGCGACGAGGCACTGTCTTCCACACCTGTCTGTGGGCT 1260
    |||
DB 1201 ATTTGTCAGCTTCCCGCGACGAGGCACTGTCTTCCACACCTGTCTGTGGGCT 1260
QY 1261 GTCCGGGAAGACACTAACCCTGTGGGCACTCAAGGTTGCGCCAGCCGCGCTGAG 1320
    |||
DB 1261 GTCCGGGAAGACACTAACCCTGTGGGCACTCAAGGTTGCGCCAGCCGCGCTGAG 1320
QY 1321 GGAGAACCTCAGGAGTTCTGTGAAGTCTCTGCAATGGAACCTCAGGCCCTGCCAACG 1380
    |||
DB 1321 GGAGAACCTCAGGAGTTCTGTGAAGTCTCTGCAATGGAACCTCAGGCCCTGCCAACG 1380
QY 1381 GACCTGAGCGTAGAGGCGCTTGTGCTCCGCAATTAATTAATTAACAAGACGCTGGGTC 1440
    |||
DB 1381 GACCTGAGCGTAGAGGCGCTTGTGCTCCGCAATTAATTAATTAACAAGACGCTGGGTC 1440
QY 1441 CGTCCATATCCAGAGAGGACAGCTGGGCCCTGCTGGAGGCTGAACCCCTGGCGAG 1500
    |||
DB 1441 CGTCCATATCCAGAGAGGACAGCTGGGCCCTGCTGGAGGCTGAACCCCTGGCGAG 1500
QY 1501 CGGATGATCGGCTTTGCTGACCGACCAAGCATGATTCGACCAAGGGCACCCCTTCAGA 1560
    |||
DB 1501 CGGATGATCGGCTTTGCTGACCGACCAAGCATGATTCGACCAAGGGCACCCCTTCAGA 1560
QY 1561 GCTAGGAGCGAAGACCGGTGTGCTCCACTCAGGGCTTGTGCTTGGGCTTCACTTG 1620
    |||
DB 1561 GCTAGGAGCGAAGACCGGTGTGCTCCACTCAGGGCTTGTGCTTGGGCTTCACTTG 1620
QY 1621 GAAAGTTCTAGTTCCCTCCAGGCTTCTAAGAACATCTGGGCCAGGGCTATAGCTGAT 1680
    |||
DB 1621 GAAAGTTCTAGTTCCCTCCAGGCTTCTAAGAACATCTGGGCCAGGGCTATAGCTGAT 1680
QY 1681 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCGAGCTCTGTTTATTTTGGTTAAA 1740
    |||
DB 1681 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCGAGCTCTGTTTATTTTGGTTAAA 1740
QY 1741 CTTATGAAATGTATTAAGAAAGTGCAGCTCGAGAGATTCAGAGATGAACACACC 1800
    |||
DB 1741 CTTATGAAATGTATTAAGAAAGTGCAGCTCGAGAGATTCAGAGATGAACACACC 1800
QY 1801 AGACCCGAGATCAACAAGCCCAACATGCCCAGGCTTCCAGAGACCCCGAGCCCAAGAC 1860
    |||
DB 1801 AGACCCGAGATCAACAAGCCCAACATGCCCAGGCTTCCAGAGACCCCGAGCCCAAGAC 1860
QY 1861 CATGTTCTGAATTTCAAGACACCGTGAAGCTGCTGTTGATCTTTAACTCATGGAAG 1920
    |||
DB 1861 CATGTTCTGAATTTCAAGACACCGTGAAGCTGCTGTTGATCTTTAACTCATGGAAG 1920
QY 1921 ATAATACTTCAAGTTTGAATAAATGTTTCTGTGAATG 1964
```

Db 1921 ATAACTACCTTACGTTTGAAATAATGTTTCTGTGTAATG 1964

RESULT 2
ID ADC37382
X AC ADC37382 standard; DNA; 1964 BP.
X AC ADC37382;

DE 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 215.

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

OS Homo sapiens.

PN WO2003048202-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.

PR 03-DEC-2001; 2001JP-0036692.

PR 05-DEC-2001; 2001US-035829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

PA (ASAH) ASAMI KASEI KK.

PI Matsuda A, Muramatsu S;

DR WPI: 2003-505282/47.

DR P-PSDB; ADC37383.

PT New purified protein that activates nuclear factor kappa B (NF-kappaB),

PT useful for treating inflammation, autoimmune diseases, cancers,

PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or

PT ischemic disorders.

PS Claim 4; SEQ ID NO 215; 938bp; English.

CC The present invention relates to novel proteins and their coding

CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-

CC kappaB). The proteins and their coding sequences are useful for treating

CC a disease associated with NF-kappaB activation, such as inflammation,

CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,

CC neurodegenerative diseases, or ischemic disorders.

SQ Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 1964; DB 10; Length 1964;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	181	ACTGCGCGGAGGCGGAGCAATTCGCGCGGAGATACGCCGAAGCAAGTGGTGGAT	240
Qy	241	ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGACGCTACAG	300
Db	241	ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGACGCTACAG	300
Qy	301	GGAATGCCATGAACATCCGGGGCCCGATGTGTCTCTCTGAAACATTGGGAAATG	360
Db	301	GGAATGCCATGAACATCCGGGGCCCGATGTGTCTCTCTGAAACATTGGGAAATG	360
Qy	361	AAAGTGAAGAAACCCCGGAGATACAGATCATTAAGAGAGGCAAGAGTCTATGAG	420
Db	361	AAAGTGAAGAAACCCCGGAGATACAGATCATTAAGAGAGGCAAGAGTCTATGAG	420
Qy	421	CACATCAGCGCATGACCGGAGCGTAAGCGGAGCATTAAGAGAGATATTTCTCAG	480
Db	421	CACATCAGCGCATGACCGGAGCGTAAGCGGAGCATTAAGAGAGATATTTCTCAG	480
Qy	481	GATCGATAGGAACCAAGAGCGGAACTTACCTCCATCTCTGCAATATGAGAGTAT	540
Db	481	GATCGATAGGAACCAAGAGCGGAACTTACCTCCATCTCTGCAATATGAGAGTAT	540
Qy	541	AAACCGAGAGTGGGCTTACTGACAGGACCTGAGCAGATGCGCGCTTGTCTCTAT	600
Db	541	AAACCGAGAGTGGGCTTACTGACAGGACCTGAGCAGATGCGCGCTTGTCTCTAT	600
Qy	601	CTTCTCTGAGAGAGATGATCTTGGGCACTGTGACAGCTGTGGCAGTAGAGGACCTCC	660
Db	601	CTTCTCTGAGAGAGATGATCTTGGGCACTGTGACAGCTGTGGCAGTAGAGGACCTCC	660
Qy	661	CTGCAAGGATTTCAAGCCCAATGCGGGGACCTGTCAGGGGCTCCAAACCAACGAG	720
Db	661	CTGCAAGGATTTCAAGCCCAATGCGGGGACCTGTCAGGGGCTCCAAACCAACGAG	720
Qy	721	CATGTGTAGCCACGCTCAACCAACCATGAGGAGCATCAGAGCAAGAAAGTCTATGT	780
Db	721	CATGTGTAGCCACGCTCAACCAACCATGAGGAGCATCAGAGCAAGAAAGTCTATGT	780
Qy	781	GCGCAGTGTTCCTCGTTAGGCTGCTCATCGGATATTTGATGACGGGATCTCTCGGG	840
Db	781	GCGCAGTGTTCCTCGTTAGGCTGCTCATCGGATATTTGATGACGGGATCTCTCGGG	840
Qy	841	CTCACCCTGCGCTGTGGAGCGTGTATCTGTAGAGGGAACAGCGCTGTATGCGGATA	900
Db	841	CTCACCCTGCGCTGTGGAGCGTGTATCTGTAGAGGGAACAGCGCTGTATGCGGATA	900
Qy	901	ACAAGATGCGCTTTAAGGTTTACGAGAGAAAGCGCTCAAGAAAGCTCCAGTGTGGCCG	960
Db	901	ACAAGATGCGCTTTAAGGTTTACGAGAGAAAGCGCTCAAGAAAGCTCCAGTGTGGCCG	960
Qy	961	TGGGACAGTTTGTGAAACCGGTTCTGTATACCTGGGCGGAGATGAGGACATGTGCTC	1020
Db	961	TGGGACAGTTTGTGAAACCGGTTCTGTATACCTGGGCGGAGATGAGGACATGTGCTC	1020
Qy	1021	AAGCATTTAAGGCTCTTATGAAAGAACTTAACAAGAAAGCAAGGGGACCTGCCACCCCA	1080
Db	1021	AAGCATTTAAGGCTCTTATGAAAGAACTTAACAAGAAAGCAAGGGGACCTGCCACCCCA	1080
Qy	1081	GCCAAACCGAGCAGAGGTCGTGCGCATTCAGGCGCTGTGCGGCTTACGTGCGGGAAG	1140
Db	1081	GCCAAACCGAGCAGAGGTCGTGCGCATTCAGGCGCTGTGCGGCTTACGTGCGGGAAG	1140
Qy	1141	ACCCTCTGCAAGGGGAGCAGGAGGCCCTCTCAGGCCCAACAGCCCGGTTCCCGGCGCC	1200
Db	1141	ACCCTCTGCAAGGGGAGCAGGAGGCCCTCTCAGGCCCAACAGCCCGGTTCCCGGCGCC	1200
Qy	1201	ATTGTGTAGCTTCCCGGCAAGGGGACCTGTTCTTCAACACCTGTCTGTGTGGGGCT	1260
Db	1201	ATTGTGTAGCTTCCCGGCAAGGGGACCTGTTCTTCAACACCTGTCTGTGTGGGGCT	1260
Qy	1261	GTCGGGAAAGACACTACCTGTGGGACCTCAGGGTGTGCCAGGCCGCGCTGACTGAG	1320
Db	1261	GTCGGGAAAGACACTACCTGTGGGACCTCAGGGTGTGCCAGGCCGCGCTGACTGAG	1320

Db 1261 GTCGGGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTTGACTCAG 1320
Qy 1321 GGAGAGCTCAGGGTCTCTGAGATTCCTGCAAGTGAATCCAGTCCCGCTCCCAAGC 1380
Db 1321 GGAGAGCTCAGGGTCTCTGAGATTCCTGCAAGTGAATCCAGTCCCGCTCCCAAGC 1380
Qy 1381 GACCTGGACGTAGAGGCGCTTGGTTCGGCCATTATGATTTAGACAGAGCTGTGGGTC 1440
Db 1381 GACCTGGACGTAGAGGCGCTTGGTTCGGCCATTATGATTTAGACAGAGCTGTGGGTC 1440
Qy 1441 CGTGGCATATCCAGAGAGACAGCTGGCCCTCTGTGGAGGCTGAACACCTGTGGAG 1500
Db 1441 CGTGGCATATCCAGAGAGACAGCTGGCCCTCTGTGGAGGCTGAACACCTGTGGAG 1500
Qy 1501 CGGGTGATCGGCTTGGCTGACCCAGACATGATTCGACAGAGGACCCCTTGACA 1560
Db 1501 CGGGTGATCGGCTTGGCTGACCCAGACATGATTCGACAGAGGACCCCTTGACA 1560
Qy 1561 GCTAGGAGCAGACAGCGGTGTCTCCACCTCAGAGGCTTGCCTGTGCGGCTTCACTTG 1620
Db 1561 GCTAGGAGCAGACAGCGGTGTCTCCACCTCAGAGGCTTGCCTGTGCGGCTTCACTTG 1620
Qy 1621 GAAAGTTCTAGTTCCTCCAGGCTTTAGAGCATGTGGCCAGGGCTCATGTGGAT 1680
Db 1621 GAAAGTTCTAGTTCCTCCAGGCTTTAGAGCATGTGGCCAGGGCTCATGTGGAT 1680
Qy 1681 AATTTCCCTAGGCTTAACAACCCAGAGAGCTTGGCGCTTATTTATTTGGTTAAA 1740
Db 1681 AATTTCCCTAGGCTTAACAACCCAGAGAGCTTGGCGCTTATTTATTTGGTTAAA 1740
Qy 1741 CTTATGAAAATGTATTAAGAAAGAGTCAGCTCAGAGAGATTCAAGATGGAACACACC 1800
Db 1741 CTTATGAAAATGTATTAAGAAAGAGTCAGCTCAGAGAGATTCAAGATGGAACACACC 1800
Qy 1801 AGAACCCAGATCACAAGCCCAATGCGGCTCCAGACACCCCGAGCCCTCACTGAC 1860
Db 1801 AGAACCCAGATCACAAGCCCAATGCGGCTCCAGACACCCCGAGCCCTCACTGAC 1860
Qy 1861 CATCGTCTGAATTCAGACGACCGAGGCTGCTTGTGATCTTAACTCATGGAAG 1920
Db 1861 CATCGTCTGAATTCAGACGACCGAGGCTGCTTGTGATCTTAACTCATGGAAG 1920
Qy 1921 ATAACCTACTTCAAGTTTGAATAATGTTTCCGTGTAATG 1964
Db 1921 ATAACCTACTTCAAGTTTGAATAATGTTTCCGTGTAATG 1964

RESULT 3
ADC37384
ID ADC37384 standard; DNA; 1964 BP.
AC ADC37384;
XX
XX 18-DEC-2003 (first entry)
DE Nuclear factor kappa B (NF-kappab) activating gene, SEQ ID 217.
XX
XX Nuclear factor kappa B, NF-kappab; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Anti-inflammatory;
KW immunomodulator; cytotoxic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Neotropic; Cardiant; Gene therapy; human; gene; de.
XX
OS Homo sapiens.
XX
XX PN MO2003048202-A2.
XX
XX 12-JUN-2003.
PD
XX
XX 03-DEC-2002; 2002WO-JP012644.
PF
XX
XX 03-DEC-2001; 2001JP-00368692.
PR
XX 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
XX (ASAH) ASAH KASEI KK.
XX
XX Matsuda A, Muramatsu S;
PI
XX
XX WPI; 2003-505282/47.
DR P-PSDB; ADC37385.
XX
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappab),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischaemic disorders.
XX
XX
PS Claim 4; SEQ ID NO 217; 938pp; English.
XX
XX The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappab). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappab activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
XX Sequence 1964 BP; 473 A; 552 C; 561 G; 378 T; 0 U; 0 Other;
SQ
Query Match 99.9%; Score 1962.4; DB 10; Length 1964;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1963; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGCGTGTAGAGGTCCGGGCAAGTGTGGGCAAGAGCGAGAGCATATTATG 60
Db 1 ATGAGCGTGTAGAGGTCCGGGCAAGTGTGGGCAAGAGCGAGAGCATATTATG 60
Qy 61 AATTAAGAAAAGGACACCGAGCTGGGCTCCAGAGACAAAGGGCTTAAGCCTTTTCCA 120
Db 61 AATTAAGAAAAGGACACCGAGCTGGGCTCCAGAGACAAAGGGCTTAAGCCTTTTCCA 120
Qy 121 AGCTACAAACAACGCTGATCATTTGGGGAATTGTACATGAGAGAGCTGCTCTCTG 180
Db 121 AGCTACAAACAACGCTGATCATTTGGGGAATTGTACATGAGAGAGCTGCTCTCTG 180
Qy 181 ACTGCCGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240
Db 181 ACTGCCGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240
Qy 241 ATGCTGGAGACTGGGAGAAATTCAAAAGCAGCAAGAAAGTCAATATGAGGTTCAAG 300
Db 241 ATGCTGGAGACTGGGAGAAATTCAAAAGCAGCAAGAAAGTCAATATGAGGTTCAAG 300
Qy 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTGTCAGTCCCTCTGAACATTGAGAAATG 360
Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTGTCAGTCCCTCTGAACATTGAGAAATG 360
Qy 361 AAGTTGAAAAACCCCGGAAGATACAGATCATGAGAGAGAGGCAAGAGTCATTTGAG 420
Db 361 AAGTTGAAAAACCCCGGAAGATACAGATCATGAGAGAGAGGCAAGAGTCATTTGAG 420
Qy 421 CAATCCAGGCACTGACCGGGACGTAAGCGGGAATTAAGAAAGCATATATCTTCAAG 480
Db 421 CAATCCAGGCACTGACCGGGACGTAAGCGGGAATTAAGAAAGCATATATCTTCAAG 480
Qy 481 GATGATACGGAACCAAGCAGCGGGAATTAATCAATCACTCTCTGGCATATGAGAGTAT 540
Db 481 GATGATACGGAACCAAGCAGCGGGAATTAATCAATCACTCTCTGGCATATGAGAGTAT 540
Qy 541 AACCCGAGGTGGCTATCTGACAGGACCTGAGCAGATGCGCGCTTGTCTCTCTAT 600
Db 541 AACCCGAGGTGGCTATCTGACAGGACCTGAGCAGATGCGCGCTTGTCTCTCTAT 600
Qy 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTCACTGTGCGCAGTGAAGCACTCC 660
Db 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTCACTGTGCGCAGTGAAGCACTCC 660

OY	6651	CTGCAGGGATTTTCAAGCCCAAAATGGGGGGAACCGTCCAGGGGCTCCAGACCAAGAG	720
Db	6651	CTGCAGGGATTTTCAAGCCCAAAATGGGGGGAACCGTCCAGGGGCTCCAGACCAAGAG	720
OY	721	CATGTGTAGCCACGTCAACACCAGACCATGTGGGGCATCAGGACAAAGAAAGATCTATGT	780
Db	721	CATGTGTAGCCACGTCAACACCAGACCATGTGGGGCATCAGGACAAAGAAAGATCTATGT	780
OY	781	GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGGATATGATTGACGGGATCTCTCCGG	840
Db	781	GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGGATATGATTGACGGGATCTCTCCGG	840
OY	841	CTCAACCTGCGGCTGTGGGACGTGTATCTGTAGAGAAAGCGAACAGGCGTGTATGCCATA	900
Db	841	CTCAACCTGCGGCTGTGGGACGTGTATCTGTAGAGAAAGCGAACAGGCGTGTATGCCATA	900
OY	901	ACAGAAATGCGCTTTAAGGTTCAAGAAAGCGCTTCAAGAGACGTCAAGGTGTGGCCG	960
Db	901	ACAGAAATGCGCTTTAAGGTTCAAGAAAGCGCTTCAAGAGACGTCAAGGTGTGGCCG	960
OY	961	TGGGACGTTTTTGGACCGGTTGCTGTATACCTGGGCGAGGGATGAGAGACATGTGCTC	1020
Db	961	TGGGACGTTTTTGGACCGGTTGCTGTATACCTGGGCGAGGGATGAGAGACATGTGCTC	1020
OY	1021	AAGCATCTTAAAGGCTCTTATGAAAGAACTAACAGAAAGAGGGGGAACCTGCACCTCCA	1080
Db	1021	AAGCATCTTAAAGGCTCTTATGAAAGAACTAACAGAAAGAGGGGGAACCTGCACCTCCA	1080
OY	1081	GCCAAACCCGAGACAGGGTCTGTGGCATTCAGGGCTGTGGCGGCTTAAAGTGGCGGAG	1140
Db	1081	GCCAAACCCGAGACAGGGTCTGTGGCATTCAGGGCTGTGGCGGCTTAAAGTGGCGGAG	1140
OY	1141	ACCCCTGTCAAGGGGGGAGAGGAGCGGCCCTCCAGGCCCAACAGCCGGTTCCCGGGGCC	1200
Db	1141	ACCCCTGTCAAGGGGGGAGAGGAGCGGCCCTCCAGGCCCAACAGCCGGTTCCCGGGGCC	1200
OY	1201	ATTGTGTCAGCTTCCCGGCACCGGACCTGTGTTTTCACAACCTGTCTGTGGGGCT	1260
Db	1201	ATTGTGTCAGCTTCCCGGCACCGGACCTGTGTTTTCACAACCTGTCTGTGGGGCT	1260
OY	1261	GTCGCGGAGAAACAACCTACCTGCTGGGGCATCAGGGGTGGCCAGCGCGGCTGTGGCTAG	1320
Db	1261	GTCGCGGAGAAACAACCTACCTGCTGGGGCATCAGGGGTGGCCAGCGCGGCTGTGGCTAG	1320
OY	1321	GAAGAGCTCAGGGTTCCTGAGAACTCTGCAAGTGGAACTCCATGCCCCGCTCCAAAG	1380
Db	1321	GAAGAGCTCAGGGTTCCTGAGAACTCTGCAAGTGGAACTCCATGCCCCGCTCCAAAG	1380
OY	1381	GAACCTGAGCTGAGAGGGGCTTGTGTCCGCCATTATGATTTTCAGACAGAGCTGCTGGGTC	1440
Db	1381	GAACCTGAGCTGAGAGGGGCTTGTGTCCGCCATTATGATTTTCAGACAGAGCTGCTGGGTC	1440
OY	1441	CGTGCATATCCAGAGAGGACCAAGCTGAGCCCCCTGTGGAGGCTGAACAACCTTGGCGAG	1500
Db	1441	CGTGCATATCCAGAGAGGACCAAGCTGAGCCCCCTGTGGAGGCTGAACAACCTTGGCGAG	1500
OY	1501	CGGGTAGATGCGCTTTCGCTGCACCCAGACCTGATTCGACCAAGGGACCCCTTCAGA	1560
Db	1501	CGGGTAGATGCGCTTTCGCTGCACCCAGACCTGATTCGACCAAGGGACCCCTTCAGA	1560
OY	1561	GCTAGGGAGCAACAGCGGTGTGGCTCCCACTCAGGGCCTTGGCCTGTGGGGCTCACTTG	1620
Db	1561	GCTAGGGAGCAACAGCGGTGTGGCTCCCACTCAGGGCCTTGGCCTGTGGGGCTCACTTG	1620
OY	1621	GAAGATTTCTAGTTCCCTCCAGGGCTTCTAGAGAGCATCTGGGCGAGGGCTCATGGCTGAT	1680
Db	1621	GAAGATTTCTAGTTCCCTCCAGGGCTTCTAGAGAGCATCTGGGCGAGGGCTCATGGCTGAT	1680
OY	1681	AATTTCCCTAAGCTTAAACACCAAGCAAGCTTGCGGCTCTCGTTTATTTTGGTTAAA	1740
Db	1681	AATTTCCCTAAGCTTAAACACCAAGCAAGCTTGCGGCTCTCGTTTATTTTGGTTAAA	1740

QY	1741	CTTATGAAAAATGTTTAAGAAAAGATCAGCTCGAGAGAGATTTCAGATGAAACACACC	1800
DB	1741	CTTATGAAAAATGTTTAAGAAAAGATCAGCTCGAGAGAGATTTCAGATGAAACACACC	1800
QY	1801	AGACCCCAAGATCACAAAGCCCAACCCACCCCTCCAGACACCCCCACACGAC	1860
DB	1801	AGACCCCAAGATCACAAAGCCCAACCCACCCCTCCAGACACCCCCACACGAC	1860
QY	1861	CATGCTTCTGAATTTCTGACCAACCCGTCGAGCTTCCTTTGATCTTTAACTCATGAAAG	1920
DB	1861	CATGCTTCTGAATTTCTGACCAACCCGTCGAGCTTCCTTTGATCTTTAACTCATGAAAG	1920
QY	1921	ATAACTACCTTCACGTTTTGAAATAAATGTTTCTGTGAAATG	1964
DB	1921	ATAACTACCTTCACGTTTTGAAATAAATGTTTCTGTGAAATG	1964
RESULT 4			
ID	ABX71403	standard; cDNA; 2072 BP.	
AC	ABX71403;		
DT	14-APR-2003	(first entry)	
DE	Human cell cycle-associated cDNA from clone DKFzphes3_35p22.		
XX	Human; gene; gene therapy; vaccine; disease treatment; detection; ss.		
OS	Homo sapiens.		
PN	W0200112659-A2.		
XX	22-FEB-2001.		
XX	18-AUG-2000; 2000WO-1B001496.		
PR	18-AUG-1999; 99US-0149499P.		
PR	28-SEP-1999; 99US-0156503P.		
FA	(GEHU-) GERMAN HUMAN GENOME PROJECT.		
PI	Wiemann S;		
DR	WPI: 2001-327840/34.		
XX	P-PSDB; ABUS3234.		
PT	Nucleic acids having the sequences of clones isolated from libraries of		
XX	different human tissues, useful in recombinant DNA methodologies.		
PS	Claim 1; Page 867; 1095pp; English.		
CC	This invention describes novel polynucleotides and polypeptides isolated		
CC	from human cDNA libraries which can be used for gene therapy or in		
CC	vaccines. The polynucleotides of the invention and antibodies encoded by		
CC	them may be used in the prevention, diagnosis and treatment of diseases		
CC	associated with inappropriate polypeptide expression. The products of the		
CC	invention may also be used to identify modulators of expression and		
CC	activity and to down regulate expression and activity. The antibodies of		
CC	the invention may also be used as diagnostic agents for detecting the		
CC	presence of polypeptides in samples. This sequence encodes a polypeptide		
CC	described in the disclosure of the invention		
SO	Sequence 2072 BP; 510 A; 579 C; 586 G; 397 T; 0 U; 0 Other;		
QY	Query Match	99.24; Score 1948; DB 5; Length 2072;	
DB	Best Local Similarity	99.54; Pred. No. 0;	
	Matches 1954; Conservative	0; Mismatches 10; Indels 0; Gaps 0;	
QY	1	ATGACAGCTGGTAAAGCTCCGGGACAGTGGTGGGACAAAGCGAAGGACATCATTTATG	60
DB	99	ATGACAGCTGGTAAAGCTCCGGGACAGTGGTGGGACAAAGCGAAGGACATCATTTATG	158

QY 61 AAATACGAAAAAGGACACCGAGCTGGCTGCAGAGACAAAGGGGCTTAAAGCTTTTTCGA 120
| | | | |
Db 159 AAATACGAAAAAGGACACCGAGCTGGCTGCAGAGACAAAGGGGCTTAAAGCTTTTTCGA 218
| | | | |
QY 121 AGCTACAAACAACAGCTGCATCATTTGGGGATTGTACATGACGAGCTGCTCTCTG 180
| | | | |
Db 219 AGCTACAAACAACAGCTGCATCATTTGGGGATTGTACATGACGAGCTGCTCTCTG 278
| | | | |
QY 181 ACTGCGGGGAGGCGAAGCAAAATTCCGCGGGAGATCAGCCGAAAGAGCAAGTGGGTGAT 240
| | | | |
Db 279 ACTGCGGGGAGGCGAAGCAAAATTCCGCGGGAGATCAGCCGAAAGAGCAAGTGGGTGAT 338
| | | | |
QY 241 ATGCTGGGAGACTGGGAAATACAAAAGACGAAAGCTCATGATCGAGCTACAG 300
| | | | |
Db 339 ATGCTGGGAGACTGGGAAATACAAAAGACGAAAGCTCATGATCGAGCTACAG 398
| | | | |
QY 301 GGAATGCCATGAAATCCGGGGCCCGATGTGTCACTCTCTGAACTTTAGAAATG 360
| | | | |
Db 399 GGAATGCCATGAAATCCGGGGCCCGATGTGTCACTCTCTGAACTTTAGAAATG 458
| | | | |
QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAAGGAGGAGGAGGAGGATCTGAG 420
| | | | |
Db 459 AAGTTGAAAAACCCCGAAGATACAGATCATGAAAGGAGGAGGAGGAGGATCTGAG 518
| | | | |
QY 421 CACATCCAGCGCATCGACCGGAGAGTAAGCGGACATTAAGAAAGCATATATTTTCAGG 480
| | | | |
Db 519 CACATCCAGCGCATCGACCGGAGAGTAAGCGGACATTAAGAAAGCATATATTTTCAGG 578
| | | | |
QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTCTCACTCTCTGAGATATGAGAGATAT 540
| | | | |
Db 579 GATCGATACGGAACCAAGCAGCGGGAATCTCTCACTCTCTGAGATATGAGAGATAT 638
| | | | |
QY 541 AACCCGGAGGTGGGCTACTGAGGAGCTGAGCCAGATCGCCGCTTTCTCTCTAT 600
| | | | |
Db 639 AACCCGGAGGTGGGCTACTGAGGAGCTGAGCCAGATCGCCGCTTTCTCTCTAT 698
| | | | |
QY 601 CTTCCTGAGAGAGATGCAATTCGTGGCACTGTGACAGCTGTGGCAGTGAAGCACTCC 660
| | | | |
Db 699 CTTCCTGAGAGAGATGCAATTCGTGGCACTGTGACAGCTGTGGCAGTGAAGCACTCC 758
| | | | |
QY 661 CTGCGAGGATTTTCAACAGCCCAATGCGGGAACCTTCAAGGGGCTCCAAAGCAACAGAG 720
| | | | |
Db 759 CTGCGAGGATTTTCAACAGCCCAATGCGGGAACCTTCAAGGGGCTCCAAAGCAACAGAG 818
| | | | |
QY 721 CATGTGTAGCAAGTGCACAAACCAAGACCATGGGGGATCAGGACAAAGAAAGATCTATGT 780
| | | | |
Db 819 CATGTGTAGCAAGTGCACAAACCAAGACCATGGGGGATCAGGACAAAGAAAGATCTATGT 878
| | | | |
QY 781 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGATATTTGATTGAACGGGATCTCTCTGGG 840
| | | | |
Db 879 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGATATTTGATTGAACGGGATCTCTCTGGG 938
| | | | |
QY 841 CTGACCTTGGCTGTGGAGCTGTATCTGGTAGAAGGCGAACAAGGCTGTGATGCCGATA 900
| | | | |
Db 939 CTGACCTTGGCTGTGGAGCTGTATCTGGTAGAAGGCGAACAAGGCTGTGATGCCGATA 998
| | | | |
QY 901 ACAAGATGCGCTTTAAGTTTCAAGCAGAAGGCGCTCAAGAAAGCTCCAGAGTGGGCGG 960
| | | | |
Db 999 ACAAGATGCGCTTTAAGTTTCAAGCAGAAGGCGCTCAAGAAAGCTCCAGAGTGGGCGG 1058
| | | | |
QY 961 TGGGCAAGTTTTCGACACCGGTTCTTGAATACCTGGGCGAGGATGAGAGACCTGTGCTC 1020
| | | | |
Db 1059 TGGGCAAGTTTTCGACACCGGTTCTTGAATACCTGGGCGAGGATGAGAGACCTGTGCTC 1118
| | | | |
QY 1021 AAGCATTTTAGGCTCTTATGAAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
| | | | |
Db 1119 AAGCATTTTAGGCTCTTATGAAAGAACTAAACAAGAAAGAGGGGACCTGCCACCCCA 1178
| | | | |
QY 1081 GCCAAACCCGAGCAGGGGTGTGGGCAATCAGGCTGTGGCGGCTTACGTTGGGAGGAG 1140
| | | | |
Db 1179 GCCAAACCCGAGCAGGGGTGTGGGCAATCAGGCTGTGGCGGCTTACGTTGGGAGGAG 1238
| | | | |
QY 1141 ACCCTCTGCAAGGGGAGCAGGACAGGCCCTTCAAGGCCCAACAGCCCGGTTCCGCGGCC 1200
| | | | |

Db 1239 ACCCTCTGCAAGGGGAGCAGGACAGGCCCTTCAAGGCCCAACAGCCCGGTTCCGCGGCC 1298
| | | | |
QY 1201 ATTTGGTCAAGCTTCCCGGACAGGGGACCTCGTTCTTCCACACCTGTCCTGGTGGGGCT 1260
| | | | |
Db 1299 ATTTGGTCAAGCTTCCCGGACAGGGGACCTCGTTCTTCCACACCTGTCCTGGTGGGGCT 1358
| | | | |
QY 1261 GTCGAGGAAGACACTACCTGAGGCACTCAGGGTGTGCCAGCCCGGCTGGGCTCAG 1320
| | | | |
Db 1359 GTCGAGGAAGACACTACCTGAGGCACTCAGGGTGTGCCAGCCCGGCTGGGCTCAG 1418
| | | | |
QY 1321 GAGAGACTTCAGGGTTCCTGAGATTCTGCACTGGAATCCATGCCCCGCTCCACAG 1380
| | | | |
Db 1419 GAGAGACTTCAGGGTTCCTGAGATTCTGCACTGGAATCCATGCCCCGCTCCACAG 1478
| | | | |
QY 1381 GACCTGAGAGTGAAGGGCCCTTGGTCCGCAATTAATGATTTTTCAGACAGAGCTGTGGTTC 1440
| | | | |
Db 1479 GACCTGAGAGTGAAGGGCCCTTGGTCCGCAATTAATGATTTTTCAGACAGAGCTGTGGTTC 1538
| | | | |
QY 1441 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGCTGGCAGGCTGAACAACCTGCGGAG 1500
| | | | |
Db 1539 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGCTGGCAGGCTGAACAACCTGCGGAG 1598
| | | | |
QY 1501 CGGCTGAGATCGGCTTTCCGTGCAACCAAGCATGATTTCCGACAGGGCACCCCTTTCAGA 1560
| | | | |
Db 1599 CGGCTGAGATCGGCTTTCCGTGCAACCAAGCATGATTTCCGACAGGGCACCCCTTTCAGA 1658
| | | | |
QY 1561 GCTAAGGAGAGAACAGCGGTGTGCTCCACTCAAGGGGCTTGGCTGCGGGCTCCACTTG 1620
| | | | |
Db 1659 GCTAAGGAGAGAACAGCGGTGTGCTCCACTCAAGGGGCTTGGCTGCGGGCTCCACTTG 1718
| | | | |
QY 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTAAGAGCATGTGGGCGAGGCTCATGAGCTGAT 1680
| | | | |
Db 1719 GAAAGTTCTCAGTTCCTCCAGGCTTCTAAGAGCATGTGGGCGAGGCTCATGAGCTGAT 1778
| | | | |
QY 1681 AATTTCCCTAAGCTTAAACAACCAAGCAAGCTTCCGCTGCTGTTTATTTTGGTTAA 1740
| | | | |
Db 1779 AATTTCCCTAAGCTTAAACAACCAAGCAAGCTTCCGCTGCTGTTTATTTTGGTTAA 1838
| | | | |
QY 1741 CTTATGAAATGTATTAAGAAAGAGAGAGCTGAGAGAGATTCAAGATGGAACAACAC 1800
| | | | |
Db 1839 CTTATGAAATGTATTAAGAAAGAGAGAGCTGAGAGAGATTCAAGATGGAACAACAC 1898
| | | | |
QY 1801 AGACCCCAATCAACAAGCCCAATGAGCCAGGCTTCCAGACACCCAGGCCCAAGAC 1860
| | | | |
Db 1899 AGACCCCAATCAACAAGCCCAATGAGCCAGGCTTCCAGACACCCAGGCCCAAGAC 1958
| | | | |
QY 1861 CATGTTCTGAATTTGACGACACCGTGAAGCTGCTTTGTACTTTAACTCATGGAAG 1920
| | | | |
Db 1959 CATGTTCTGAATTTGACGACACCGTGAAGCTGCTTTGTACTTTAACTCATGGAAG 2018
| | | | |
QY 1921 ATTAACAACCTTCAAGTTTGAATTAATGTTTCTGTGAAATG 1964
| | | | |
Db 2019 ATTAACAACCTTCAAGTTTGAATTAATGTTTCTGTGAAATG 2062
| | | | |

RESULT 5
AD084596
ID AD084596 standard; cDNA; 2064 BP.
XX
AC AD084596;
XX
DT 07-OCT-2004 (first entry)
XX
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1410.
XX
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN MO2004060270-A2.
XX

PD 22-JUL-2004.
XX
XX 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-0418988P.
XX
XX (GENTH) GENENTECH INC.
XX (MUTD/) MU T D.
XX (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;
XX
XX WPI, 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 1410; 5504bp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT CDNA sequence from the present invention.
XX
XX Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;
XX
XX Query Match 99.0%; Score 1944; DB 13; Length 2064;
XX Best Local Similarity 99.6%; Pred. No. 0;
XX Matches 1959; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

DB 219 AGCTACAAACAACACTCCATCATTTGGGGATTGTACAAGTAGAGGAGCTCCCTC 278
|||
QY 179 TGACTGCGCGGAGCGGAAGCAATTCGCGGAGATCAAGCCGAAGCAATGAGTGG 238
|||
DB 279 TGACTGCGCGGAGCGGAAGCAATTCGCGGAGATCAAGCCGAAGCAATGAGTGG 338
|||
QY 239 ATATGCTGGGAGACTGGGAGAAATACAAAAGCAGAAAGCTCATAGATCCAGCTACA 298
|||
DB 339 ATATGCTGGGAGACTGGGAGAAATACAAAAGCAGAAAGCTCATAGATCCAGCTACA 398
|||
QY 299 AGGAATGCCCATGAAACATCCGGGGCCGATGTCAGTCCCTCTGAACATTGAGAAA 358
|||
DB 399 AGGAATGCCCATGAAACATCCGGGGCCGATGTCAGTCCCTCTGAACATTGAGAAA 458
|||
QY 359 TGAAGTTGAAAAAACCAGGAAGATACAGATCAAGAGAGAAAGGCAAGGTCACTTG 418
|||
DB 459 TGAAGTTGAAAAAACCAGGAAGATACAGATCAAGAGAGAAAGGCAAGGTCACTTG 518
|||
QY 419 AGCAATTCAGCGCATTCAGCCGGAGCTAAGCGGAGCATTTAGAGAACATATTTCTCA 478
|||
DB 519 AGCAATTCAGCGCATTCAGCCGGAGCTAAGCGGAGCATTTAGAGAACATATTTCTCA 578
|||
QY 479 GGGATGATACGGAACCAAGCAGCGGGAATCACTCCATCCCTCGGATATAGAGAGT 538
|||
DB 579 GGGATGATACGGAACCAAGCAGCGGGAATCACTCCATCCCTCGGATATAGAGAGT 638
|||
QY 539 ATAAACCCGAGGTGGGCTACTGACAGGACCTGACCAATCGCCCTTGTCTCTCT 598
|||
DB 639 ATAAACCCGAGGTGGGCTACTGACAGGACCTGACCAATCGCCCTTGTCTCTCT 698
|||
QY 599 ATCTTCTTGAAGAGATGATCTTGTGGCACTGTGTGACCTGTGCTGCAATGAGCACT 658
|||
DB 699 ATCTTCTTGAAGAGATGATCTTGTGGCACTGTGTGACCTGTGCTGCAATGAGCACT 758
|||
QY 659 CCTGCAAGGATTTTCAACAGCCCAAAATGCGGAGCGCTCAGGGGCTTCAAGCAAG 718
|||
DB 759 CCTGCAAGGATTTTCAACAGCCCAAAATGCGGAGCGCTCAGGGGCTTCAAGCAAG 818
|||
QY 719 AGCATGTGTAGCAACGTCACAAACCAAGACATGAGGAGATGAGCAAGAAAGATCTAT 778
|||
DB 819 AGCATGTGTAGCAACGTCACAAACCAAGACATGAGGAGATGAGCAAGAAAGATCTAT 878
|||
QY 779 GTGGGCAAGTTCCTCCGTTAGGCTGCTATCCGATATTTGACGGGATCTCTCTG 838
|||
DB 879 GTGGGCAAGTTCCTCCGTTAGGCTGCTATCCGATATTTGACGGGATCTCTCTG 938
|||
QY 839 GGCTCAACCTGCGCTGTGGGAGCTGTATCTGTGAAGAGCAAGGGCTTGAATGCGGA 898
|||
DB 939 GGCTCAACCTGCGCTGTGGGAGCTGTATCTGTGAAGAGCAAGGGCTTGAATGCGGA 998
|||
QY 899 TAACAAGATTCGCTTTAAGTTTACAGAAAGCGCTCAGAAAGCTCAGATGTGGCC 958
|||
DB 999 TAACAAGATTCGCTTTAAGTTTACAGAAAGCGCTCAGAAAGCTCAGATGTGGCC 1058
|||
QY 959 CGTGGGCACTTTTTCGAAACCGGTTCTGTGATCTGTGGCCAGGAGTGGGACACTGTGC 1018
|||
DB 1059 CGTGGGCACTTTTTCGAAACCGGTTCTGTGATCTGTGGCCAGGAGTGGGACACTGTGC 1118
|||
QY 1019 TCAGCATCTTGAAGGCTCTATAGAAAGAACTAACAAGAAAGAGGGGAGCTGCCACCC 1078
|||
DB 1119 TCAGCATCTTGAAGGCTCTATAGAAAGAACTAACAAGAAAGAGGGGAGCTGCCACCC 1178
|||
QY 1079 CAGCAAAACCCAGAGCAAGGCTGTGCGGATCCAGGCTGTGCGGCTTCACTGTGGGGA 1138
|||
DB 1179 CAGCAAAACCCAGAGCAAGGCTGTGCGGATCCAGGCTGTGCGGCTTCACTGTGGGGA 1238
|||
QY 1139 AGACCTCTGCAAGGGGAGCAGGCGGCTCTCAGGCTCAGACCAAGCCGATTCGCGGCG 1198
|||
DB 1239 AGACCTCTGCAAGGGGAGCAGGCGGCTCTCAGGCTCAGACCAAGCCGATTCGCGGCG 1298
|||
QY 1199 CCAATTTGTCAGCTTCCCGGAGGAGGAGCTGTTCTTCCACACCTGTCTGTGGGG 1258
|||


```
Db 1299 CCATTGGTACGCTTCCCCCGACGCGGACACCTGCTTCCACACCTGTCCTGTGGG 1358
Qy 1259 CTGTCGGGGAAGACCTACCTGCTGGGACCTAGGGGTGGCCAGCCCGGCTGTGCTC 1318
Db 1359 CTGTCGGGGAAGACCTACCTGCTGGGACCTAGGGGTGGCCAGCCCGGCTGTGCTC 1418
Qy 1319 AGGAGAGACCTCAGGGTTCCTGAGATTCTGACAGTGAAGTCCATGCCGCTCCCAA 1378
Db 1419 AGGAGAGACCTCAGGGTTCCTGAGATTCTGACAGTGAAGTCCATGCCGCTCCCAA 1478
Qy 1379 CGGACCTGAGAGTGAAGGCGCTTGTGTCGCTTATGATTGATGACAGAGTGTGGG 1438
Db 1479 CGGACCTGAGAGTGAAGGCGCTTGTGTCGCTTATGATTGATGACAGAGTGTGGG 1538
Qy 1439 TCCGTGCTATATCCAGAGAGACAGCTGCGCCCTGCTGACGCTGGAACACCTGCGG 1498
Db 1539 TCCGTGCTATATCCAGAGAGACAGCTGCGCCCTGCTGACGCTGGAACACCTGCGG 1598
Qy 1499 AGCGGTGAGATCGGCTTTCGCTGACCCAGCACTGATTCCGACAGGCGCACCCCTTCA 1558
Db 1599 AGCGGTGAGATCGGCTTTCGCTGACCCAGCACTGATTCCGACAGGCGCACCCCTTCA 1658
Qy 1559 GAGCTAGGAGCAAGAACAGCCGTGCTCCCACTCAGGCGCTTGTGCTGCGGCTCCACT 1618
Db 1659 GAGCTAGGAGCAAGAACAGCCGTGCTCCCACTCAGGCGCTTGTGCTGCGGCTCCACT 1718
Qy 1619 TGGAAATTTCTCAGTTCCTCCAGGCTTCTAGAGAGCATGCGGCGAGGCGCTCATGCGTGG 1678
Db 1719 TGGAAATTTCTCAGTTCCTCCAGGCTTCTAGAGAGCATGCGGCGAGGCGCTCATGCGTGG 1778
Qy 1679 ATATATTCCTAGGCTTACACCAAGCAAGCTTCGCGTCTGCTTATATTTTGGTTA 1738
Db 1779 ATATATTCCTAGGCTTACACCAAGCAAGCTTCGCGTCTGCTTATATTTTGGTTA 1838
Qy 1739 AACTTATGAAATGATTAAGAAAGAGTGCAGCTGAGAGAGATTGAGAGTGAACACA 1798
Db 1839 AACTTATGAAATGATTAAGAAAGAGTGCAGCTGAGAGAGATTGAGAGTGAACACA 1898
Qy 1799 CCAAGACCCAGATCAGAAAGCAAGCCAGGCGCCCTCCAGACCCCGAGCGCAGC 1858
Db 1899 CCAAGACCCAGATCAGAAAGCAAGCCAGGCGCCCTCCAGACCCCGAGCGCAGC 1958
Qy 1859 ACCATCGTTCGATTTGACGACACCGTGAGCTGCTTGTATCTTAAACTGAGAA 1918
Db 1959 ACCATCGTTCGATTTGACGACACCGTGAGCTGCTTGTATCTTAAACTGAGAA 2018
Qy 1919 GGATTAACCTTACCTGTTGAAATTAATGTTCTGTTGAATG 1964
Db 2019 GGATTAACCTTACCTGTTGAAATTAATGTTCTGTTGAATG 2064

RESULT 6
AD086595 standard; cDNA; 2064 BP.
XX
AC AD086595;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #3468.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN MO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
```

```
XX (GETH ) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y,
DR WP1; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
PS Claim 1; SEQ ID NO 3468; 5504bp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT) a
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein in a
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;

Query Match 99.0%; Score 1944; DB 13; Length 2064;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1959; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 1 ATGAGCTGTAGAGTCCGGGACGTTGTGGGCAAGAGCGAGAGACATCATTTATG 60
Db 99 ATGAGCTGTAGAGTCCGGGACGTTGTGGGCAAGAGCGAGAGACATCATTTATG 158
Qy 61 AATTAAGAAAAGGACACCGAGCTGGGCTCCAGAGACAAAGGGGCTTAAGCTTTTGA 120
Db 159 AATTAAGAAAAGGACACCGAGCTGGGCTCCAGAGACAAAGGGGCTTAAGCTTTTGA 218
Qy 121 AGCTACAAACAACGCTGATCATTTGGGAGATTGTACA--TGAGACGAGAGCTCTCTC 178
Db 219 AGCTACAAACAACGCTGATCATTTGGGAGATTGTACAAGTAGAGAGAGCTCTCTC 278
Qy 179 TGACTGCGCGGAGGCGAAGCAATTCCGCGGAGATCAAGCCGAAAGAGCAAGTGGGTG 238
```


Db 279 TGAATGCGGGAGGCGGAACAAATTCGGCGGGAGATCAGCCGAAGAGCAAGTGGGTG 338
 QY 229 ATATGCTGGGAGATCTGGGGAATTAACAAAGCAGGAAGCTCATAGATCCAGCTTACA 298
 Db 339 ATATGCTGGGAGATCTGGGGAATTAACAAAGCAGGAAGCTCATAGATCCAGCTTACA 398
 QY 299 AGGGAATGCCATGAAATATCCGGGAGCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
 Db 399 AGGGAATGCCATGAAATATCCGGGAGCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
 QY 359 TGAAGTTGAAAAACCCCGGAAGATACAGATCATGAAAGAGAGGCAAGGTCATCTG 418
 Db 459 TGAAGTTGAAAAACCCCGGAAGATACAGATCATGAAAGAGAGGCAAGGTCATCTG 518
 QY 419 AGCAGATCCAGGCGATCCGAGCCGGAAGCTTAAGGGGACATTAAGAGAGCATATTTCTTCA 478
 Db 519 AGCAGATCCAGGCGATCCGAGCCGGAAGCTTAAGGGGACATTAAGAGAGCATATTTCTTCA 578
 QY 479 GGGATCGATACCGAAACCAAGCAGCGGGAATCTACATCTCTGCTGCTGCTGCTGCTGCT 538
 Db 579 GGGATCGATACCGAAACCAAGCAGCGGGAATCTACATCTCTGCTGCTGCTGCTGCTGCT 638
 QY 539 ATAAACCCGAGAGGTGGCTACTGACAGGACCTGAGCCACATCGCCGCTTCTCTCTCT 598
 Db 639 ATAAACCCGAGAGGTGGCTACTGACAGGACCTGAGCCACATCGCCGCTTCTCTCTCT 698
 QY 599 ATCTTCTGAGAGAGATGATCTCTGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
 Db 699 ATCTTCTGAGAGAGATGATCTCTGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
 QY 659 CCTTCAGAGGATTTTCAAGCCCAATATGGCGGAGCCGTCAGGGGCTCCAAAGCAACAGG 718
 Db 759 CCTTCAGAGGATTTTCAAGCCCAATATGGCGGAGCCGTCAGGGGCTCCAAAGCAACAGG 818
 QY 719 AGCATGTGTAGCCCAAGTCAACCCCAAGACCATGAGGACATCAGGCAAGAAAGATCTAT 778
 Db 819 AGCATGTGTAGCCCAAGTCAACCCCAAGACCATGAGGACATCAGGCAAGAAAGATCTAT 878
 QY 779 GTGGGAGATGTTCCCGTATAGGCTGCTCATATCCGATATTTGATTTGACGGGATCTCTG 838
 Db 879 GTGGGAGATGTTCCCGTATAGGCTGCTCATATCCGATATTTGATTTGACGGGATCTCTG 938
 QY 839 GGGTCAACCTGCGGCTGTGAGGAGTGTATCTGTGAGAGGCGAACAGGGGCTTATGAGCGCA 898
 Db 939 GGGTCAACCTGCGGCTGTGAGGAGTGTATCTGTGAGAGGCGAACAGGGGCTTATGAGCGCA 998
 QY 899 TAAACAAGATGCTTTAAGTTTCAAGAAAGCGCTTCAAGAGAGCTTCAAGTGTGCGC 958
 Db 999 TAAACAAGATGCTTTAAGTTTCAAGAAAGCGCTTCAAGAGAGCTTCAAGTGTGCGC 1058
 QY 959 CGTGGGACGTTTTTGGCAACCGGTTCTTATATCTTGGGCCAGGGATGAGCACTGTGTC 1018
 Db 1059 CGTGGGACGTTTTTGGCAACCGGTTCTTATATCTTGGGCCAGGGATGAGCACTGTGTC 1118
 QY 1019 TCAAGCATCTTAGGGGCTCTATGAAAGAACTTAACAAGAAAGAGGGGAGACCTGCGCAACCC 1078
 Db 1119 TCAAGCATCTTAGGGGCTCTATGAAAGAACTTAACAAGAAAGAGGGGAGACCTGCGCAACCC 1178
 QY 1079 CAGCCAAACCCGAGCAGAGGCTGTGAGCATCCAGGCTGTGCGGCTTCAAGTGTGCGGGA 1138
 Db 1179 CAGCCAAACCCGAGCAGAGGCTGTGAGCATCCAGGCTGTGCGGCTTCAAGTGTGCGGGA 1238
 QY 1139 AGACCTTCTGCAAGGGGAGCAGGAGGCTTCCAGGCTTCCAGGCTTCCGCGGCGC 1198
 Db 1239 AGACCTTCTGCAAGGGGAGCAGGAGGCTTCCAGGCTTCCAGGCTTCCGCGGCGC 1298
 QY 1199 CCAATTGGTCAAGCTTCCCGCAGAGGAGCCTGCTTCCAGCAACCTGCTGTGTTGGG 1258
 Db 1299 CCAATTGGTCAAGCTTCCCGCAGAGGAGCCTGCTTCCAGCAACCTGCTGTGTTGGG 1358
 QY 1259 CTGTCCGGGAAGACACTACCTGTGTGGCACTCAGAGGTGTGCCAGCCGCGCTGTGCTC 1318
 Db 1359 CTGTCCGGGAAGACACTACCTGTGTGGCACTCAGAGGTGTGCCAGCCGCGCTGTGCTC 1418

QY 1319 AGGAGAGACTCAGAGGTTCTTGGAGATTCCTCTCAGTGTGAACCTCATGCCCCCTCCAA 1378
 Db 1419 AGGAGAGACTCAGAGGTTCTTGGAGATTCCTCTCAGTGTGAACCTCATGCCCCCTCCAA 1478
 QY 1379 CGGACCTGAGCTAGAGGAGCCCTTGGTTCCGCAATTAATTAATTCAGACAGAGCTGTGG 1438
 Db 1479 CGGACCTGAGCTAGAGGAGCCCTTGGTTCCGCAATTAATTAATTCAGACAGAGCTGTGG 1538
 QY 1439 TCCGTGCCATATCCAGAGAGCAGCTGAGCCCTGCTGAGGCTGAACAACCTTGC 1498
 Db 1539 TCCGTGCCATATCCAGAGAGCAGCTGAGCCCTGCTGAGGCTGAACAACCTTGC 1598
 QY 1499 AGCGGCTGAGATCGGCTTTCGCTGAGACCCAGACATCTGCAACAGAGGACCCCTTCA 1558
 Db 1599 AGCGGCTGAGATCGGCTTTCGCTGAGACCCAGACATCTGCAACAGAGGACCCCTTCA 1658
 QY 1559 GAGCTAGGAGCAGACAGCCGCTGCTCCCACTCAGAGGCTTGCCTGCGGCTTCACCT 1618
 Db 1659 GAGCTAGGAGCAGACAGCCGCTGCTCCCACTCAGAGGCTTGCCTGCGGCTTCACCT 1718
 QY 1619 TGGAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAACATCTGAGCCAGGCTCATGCTGG 1678
 Db 1719 TGGAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAACATCTGAGCCAGGCTCATGCTGG 1778
 QY 1679 ATATATTTCCCTAGGCTTAAACAACCAAGACAGCTTCCGCTCTGTTTATTTTGGTTA 1738
 Db 1779 ATATATTTCCCTAGGCTTAAACAACCAAGACAGCTTCCGCTCTGTTTATTTTGGTTA 1838
 QY 1729 AACTATGAAATGTATTAAGAAAGATGACGCTGAGAGATTCAGAGATGGAACA 1798
 Db 1829 AACTATGAAATGTATTAAGAAAGATGACGCTGAGAGATTCAGAGATGGAACA 1898
 QY 1799 CCAGACCCCAAGTCAACAAAGCAACATGAGCCAGCCCTCCAGACACCCCAAGCCCAAG 1858
 Db 1899 CCAGACCCCAAGTCAACAAAGCAACATGAGCCAGCCCTCCAGACACCCCAAGCCCAAG 1958
 QY 1859 ACCATGCTTCTGATCTGAGACACCGGTGAGCTGCTGTTTAACTCATGGA 1918
 Db 1959 ACCATGCTTCTGATCTGAGACACCGGTGAGCTGCTGTTTAACTCATGGA 2018
 QY 1919 GGAATACCTACCTTCAAGTTTGAATTAATGTTTCTGTTGAATG 1964
 Db 2019 GGAATACCTTCAATGTTTGAATTAATGTTTCTGTTGAATG 2064

RESULT 7
 ID AAA15001 standard; cDNA, 1993 BP.
 XX AAA15001;
 AC 21-AUG-2000 (first entry)
 DT
 XX
 DE cDNA encoding a human proliferation and apoptosis related protein.
 XX
 KW Human; proliferation and apoptosis related protein; PROAP; psoriasis;
 KW cell proliferative disorder; immunological disorder; hepatitis;
 KW reproductive disorder; arteriosclerosis; cirrhosis; melanoma; lymphoma;
 KW cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia;
 KW aschma; diabetes mellitus; osteoarthritis; endometriosis;
 KW uterine fibroid; menstrual cycle; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 42..1691
 FT /*tag= a
 FT /product= "proliferation and apoptosis related protein"
 XX
 PN W0200023589-A2.
 XX
 PD 27-APR-2000.

XX 19-OCT-1999; 99NO-US024511.
PF 20-OCT-1998; 98US-0172216P.
PR 04-FEB-1999; 99US-0118559P.
PR 11-FEB-1999; 99US-0172229P.
PR 22-APR-1999; 99US-0154336P.
XX
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lai P;
PI Azimzai Y, Baughn MR, Yang J, Shih LL;
XX MPI: 2000-339688/29.
DR P-PSDB; MAY84901.
XX
XX New human proliferation and apoptosis related protein polypeptides used
PT for diagnosis, treatment and prevention of cell proliferative,
PT immunological and reproductive disorders.
XX
XX Claim 9; Page 117-118; 128pp; English.
XX
XX The present sequence encodes a human proliferation and apoptosis related
CC protein (PROAP). The polypeptides and polynucleotides can be used for the
CC diagnosis, treatment and prevention of cell proliferative, immunological
CC and reproductive disorders. Disorders associated with decreased
CC expression or activity of include arteriosclerosis, cirrhosis, hepatitis,
CC psoriasis, melanoma, lymphoma and cancers of the breast, brain and
CC prostate, acquired immune deficiency syndrome (AIDS), allergies, anaemia,
CC asthma, diabetes mellitus, osteoarthritis, endometriosis, uterine
CC fibroids and disruptions of the menstrual cycle. Antibodies against PROAP
CC can be used in diagnosis of disorders characterized by PROAP e.g. in ELISA
CC (enzyme linked immunosorbent assays) and the polynucleotides may be used
CC to detect and quantify gene expression in biopsied tissues. These
CC techniques can also be used to monitor regulation of PROAP levels during
CC therapeutic intervention
XX
SQ Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;
Query Match 98.3%; Score 1929.8; DB 3; Length 1993;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 ATGAGCTGGTATAGGTCGGGCGAGTGGTGGGCAAGAGGAGGACATCATTTATG 60
DB 42 ATGAGCTGGTATAGGTCGGGCGAGTGGTGGGCAAGAGGAGGACATCATTTATG 101
QY 61 AATAAGAAAAGGAGACCCGAGCTGGGCTGCCAGAGACAAAGGGCCCTTAAGCCTTTTGA 120
DB 102 AATAAGAAAAGGAGACCCGAGCTGGGCTGCCAGAGACAAAGGGCCCTTAAGCCTTTTGA 161
QY 121 AGTTACAAACAACAGTCATCTTTGGGATTTGTACATAGACGAGCTGCTCTCTG 180
DB 162 AGTTACAAACAACAGTCATCTTTGGGATTTGTACATAGACGAGCTGCTCTCTG 221
QY 181 ACTGCGGGAGGCGAAGCAATTCGGCGGAGATCAGCGGAAGAGCAAGTGGTGGAT 240
DB 222 ACTGCGGGAGGCGAAGCAATTCGGCGGAGATCAGCGGAAGAGCAAGTGGTGGAT 281
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATAGTCAGCGTACAG 300
DB 282 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATAGTCAGCGTACAG 341
QY 301 GGAATGCCCATGAACATCCGGGGCCCGATGGTCAATCTCTGTAACATTGAGAAATG 360
DB 342 GGAATGCCCATGAACATCCGGGGCCCGATGGTCAATCTCTGTAACATTGAGAAATG 401
QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGAGCAAGAGTCAATCGAG 420
DB 402 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGAGCAAGAGTCAATCGAG 461
QY 421 CACATCAGCGCATCGACCCGGAGCTTAAGCGGAGCATTAAGAGACATATATTTCTTCAGG 480

DB 462 CACATCAGCGCATCGACCCGGAGCTTAAGCGGAGCATTAAGAGACATATATTTCTTCAGG 521
QY 481 GATCGATACGGAACCAAGAGCGGGAACTACTCCATATCTCTCGCATATAGAGATAT 540
DB 522 GATCGATACGGAACCAAGAGCGGGAACTACTCCATATCTCTCGCATATAGAGATAT 581
QY 541 AACCCGGAGGTGGGCTACTGCAAGGAGCCTGAGCCACATGCGCGCTTGTCTCTCTAT 600
DB 582 AACCCGGAGGTGGGCTACTGCAAGGAGCCTGAGCCACATGCGCGCTTGTCTCTCTAT 641
QY 601 CTTCCTGAGAGAGATTCATTCGGGCACTGGGAGCTGCTGCGCAAGTAGAGGACCTCC 660
DB 642 CTTCCTGAGAGAGATTCATTCGGGCACTGGGAGCTGCTGCGCAAGTAGAGGACCTCC 701
QY 661 CTGCAAGGATTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAAGCAACAGAG 720
DB 702 CTGCAAGGATTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAAGCAACAGAG 761
QY 721 CATGTGTAGCCACGTCACAAACCAAGACATGGGGCATCAGAGACAAAGAAAGTATATGT 780
DB 762 CATGTGTAGCCACGTCACAAACCAAGACATGGGGCATCAGAGACAAAGAAAGTATATGT 821
QY 781 GGGCAATGTTCCCGTTAGGCTGCTCATCCGGATATAGATTGAGCGGATCTCTCCGGG 840
DB 822 GGGCAATGTTCCCGTTAGGCTGCTCATCCGGATATAGATTGAGCGGATCTCTCCGGG 881
QY 841 CTCACCCCTGCGCTGTGGAGCGTATCTGTATGAAGCGGAACAGCGTTGATGCCGATA 900
DB 882 CTCACCCCTGCGCTGTGGAGCGTATCTGTATGAAGCGGAACAGCGTTGATGCCGATA 941
QY 901 ACAAGATGCGCTTTAAGGTTACAGAGAGCCCTTCACAGAGACGTCAGGTGTGGCCGG 960
DB 942 ACAAGATGCGCTTTAAGGTTACAGAGAGCCCTTCACAGAGACGTCAGGTGTGGCCGG 1001
QY 961 TGGGCACTTTTTCGACACCGGTTCTGTATACCTGGGCCAGGAGTAGAGACACTGTGCTC 1020
DB 1002 TGGGCACTTTTTCGACACCGGTTCTGTATACCTGGGCCAGGAGTAGAGACACTGTGCTC 1061
QY 1021 AAGCATCTTAGGGCTCTTATGAAGAACTAACAAGAAACAGGGGACCTGCAACCCCA 1080
DB 1062 AAGCATCTTAGGGCTCTTATGAAGAACTAACAAGAAACAGGGGACCTGCAACCCCA 1121
QY 1081 GCCAAACCCGAGCAAGGATGTCGCGATCAGGCTGTGCTCGGCTTACGTGGCGGGAAG 1140
DB 1122 GCCAAACCCGAGCAAGGATGTCGCGATCAGGCTGTGCTCGGCTTACGTGGCGGGAAG 1181
QY 1141 ACCCTTGAAGGGGAGCAGGAGGCTCTCCAGGCTCCACAGCCGGTTCCCGCGGCC 1200
DB 1182 ACCCTTGAAGGGGAGCAGGAGGCTCTCCAGGCTCCACAGCCGGTTCCCGCGGCC 1241
QY 1201 ATTGGTCAAGCTTCCCGGCAAGGCACTGCTTTCACACACCTGTCTGTGGTGGGCT 1260
DB 1242 ATTGGTCAAGCTTCCCGGCAAGGCACTGCTTTCACACACCTGTCTGTGGTGGGCT 1301
QY 1261 GTTCGGGAAAGACACTTACCTGTGGCACTCAGGGTGTGCCAGCCCGGCTCTGCTCAG 1320
DB 1302 GTTCGGGAAAGACACTTACCTGTGGCACTCAGGGTGTGCCAGCCCGGCTCTGCTCAG 1361
QY 1321 GGAAGACCTCAAGGTTCTGAGATTCCTGCAAGTGAATCTCAATGCCCCCGCTCCCAAG 1380
DB 1362 GGAAGACCTCAAGGTTCTGAGATTCCTGCAAGTGAATCTCAATGCCCCCGCTCCCAAG 1421
QY 1381 GACCTGAGAGTGAAGGACCTTGGTTCCGCAATATATTTTCAAGACAGAGCTGTGGGCT 1440
DB 1422 GACCTGAGAGTGAAGGACCTTGGTTCCGCAATATATTTTCAAGACAGAGCTGTGGGCT 1481
QY 1441 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGTGTGAGGCTTGAACACCTGTGGAG 1500
DB 1482 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGTGTGAGGCTTGAACACCTGTGGAG 1541
QY 1501 CGGTTGAGATCGGCTTTCGCTGACCCAGACACTGATTTCCGACAGGAGCACCCCTTCAGA 1560
DB 1542 CGGTTGAGATCGGCTTTCGCTGACCCAGACACTGATTTCCGACAGGAGCACCCCTTCAGA 1601

QY 1561 GCTAGGAGCAAGCCGCTGCTCCCACTCAGGCGCTTGCCTTGGCGCTCCACTTG 1620
 DB 1602 GCTAGGAGCAAGCCGCTGCTCCCACTCAGGCGCTTGCCTTGGCGCTCCACTTG 1661
 QY 1621 GAAAGTTCTCAGTCTCCCTCCAGGCTTCTAGAGCACTTGGGCGGCTCATGCTGAT 1680
 DB 1662 GAAAGTTCTCAGTCTCCCTCCAGGCTTCTAGAGCACTTGGGCGGCTCATGCTGAT 1721
 QY 1681 AATTTCCCTAGGCTTAAACCAAGCAAGCTTGGCTCTCTGTTTATTTTGGTTAA 1740
 DB 1722 AATTTCCCTAGGCTTAAACCAAGCAAG-CTTGGATCTCTCTTATTTTGGTTAA 1780
 QY 1741 CTTATGAAATGATTAAGAAAGTGCAGCTGCAGAGATTCAGAGATGAAACACCC 1800
 DB 1781 CTTATGAAATGATTAAGAAAGTGCAGCTGCAGAGATTCAGAGATGAAACACCC 1840
 QY 1801 AGACCCGAGATCAAGAAAGCAACATGCCCGCTCCAGCAAGCCCGGCGGCGAGC 1860
 DB 1841 AGACCCGAGATCAAGAAAGCAACATGCCCGCTCCAGCAAGCCCGGCGGCGAGC 1900
 QY 1861 CATGCTTCTGAATTTCTGAGCAAGCCGCTGCTTGTACTTTAACTCATGGAAG 1920
 DB 1901 CATGCTTCTGAATTTCTGAGCAAGCCGCTGCTTGTACTTTAACTCATGGAAG 1960
 QY 1921 AATACTACCTTCAAGTTTGAATAATGTTTC 1953
 DB 1961 AATACTACCTTCAAGTTTGAATAATGTTTC 1993

RESULT 8

ADT88069
 ID ADT88069 standard; cDNA; 1993 BP.

AC ADT88069;

DT 30-DEC-2004 (first entry)

XX Human proliferation and apoptosis related protein (PROAP)-12 cDNA.

KM PROAP; proliferation and apoptosis related protein;
 KM cell proliferative disorder; cancer; atherosclerosis;
 KM immunological disorder; AIDS; acquired immunodeficiency syndrome;
 KM allergy; reproductive disorder; infertility; gene therapy; cytostatic;
 KM antiarteriosclerotic; immunosuppressive; anti-HIV; antiallergic;
 KM antiinfectivity; gynaecological; human; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 42..1691
 FT /*tag= a
 FT /product= "Proliferation and apoptosis related protein
 (PROAP)-12"

PN US2004203106-A1.

XX 14-OCT-2004.

PF 05-MAY-2004; 2004US-00839882.

PR 19-JAN-1999; 99US-0172216P.

PR 04-FEB-1999; 99US-0118559P.

PR 11-FEB-1999; 99US-0172229P.

PR 22-APR-1999; 99US-0154336P.

PR 19-OCT-1999; 99WO-US024511.

PR 11-APR-2001; 2001US-00807452.

XX (INCY-) INCYTE CORP.

XX Tang YF, Yue H, Hallman JL, Guejler KJ, Corley NC, Lai P;
 PI Azimzal Y, Baughn MR, Yang J, Shih LL;
 XX

DR WPI; 2004-728011/71.
 DR P-PSDB; ADT88050.
 PT New human proliferation and apoptosis related proteins and
 PT polynucleotides for diagnosing, preventing or treating disorders
 PT associated with aberrant protein expression, e.g. cancer, AIDS,
 PT atherosclerosis or infertility.
 PS Claim 5; SEQ ID NO 31; 85bp; English.
 XX
 CC The present invention relates to the human proliferation and apoptosis
 CC related protein (PROAP) and its encoding nucleic acid. The invention is
 CC useful for diagnosing, preventing or treating disorders associated with
 CC altered expression or activity of human PROAP, such as cell proliferative
 CC (e.g. cancer or atherosclerosis), immunological (e.g. acquired
 CC immunodeficiency syndrome (AIDS) or allergies) or reproductive (e.g.
 CC infertility) disorders. The invention is also be used in screening for
 CC drugs that may be used for treating or preventing the disorders mentioned
 CC above and in gene therapy. The present sequence is the human PROAP-12
 CC protein encoding cDNA.

XX Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Query Match 98.3%; Score 1929.8; DB 13; Length 1993;
 Best local Similarity 99.6%; Pred. No. 0;
 Matches 1945; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 ATGACGTGTGTAAGGTGCGGCGCAAGTTGGTGGGCAAGAGCGAGAGACATCATATG 60
 DB 42 ATGACGTGTGTAAGGTGCGGCGCAAGTTGGTGGGCAAGAGCGAGAGACATCATATG 101
 QY 61 AATATGAAAGAGGACACCGAGCTGGGCTCCAGAGACAAAGGGCTTAAGCTTTTGA 120
 DB 102 AATATGAAAGAGGACACCGAGCTGGGCTCCAGAGACAAAGGGCTTAAGCTTTTGA 161
 QY 121 AGCTACAAACAACAGTCGATCATTTTGGGATTTGTACATGAGACGAGGCTGCTCTG 180
 DB 162 AGCTACAAACAACAGTCGATCATTTTGGGATTTGTACATGAGACGAGGCTGCTCTG 221
 QY 181 ACTGCGCGGAGCGGAGCAAGCAATTCGCGGAGATCAAGCCGAAAGCAAGTGGTAT 240
 DB 222 ACTGCGCGGAGCGGAGCAAGCAATTCGCGGAGATCAAGCCGAAAGCAAGTGGTAT 281
 QY 241 ATGCTGGGAGATTTGGGAGAAATACAAAGACAGCAAGAACTCATATGAGCTTCAAG 300
 DB 282 ATGCTGGGAGATTTGGGAGAAATACAAAGACAGCAAGAACTCATATGAGCTTCAAG 341
 QY 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGATCTCTGCAACTTGAAGGAATG 360
 DB 342 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGATCTCTGCAACTTGAAGGAATG 401
 QY 361 AAGTTGAAACCCCGGAAGATACCATGATCATGAAGAGAGGAGAGTCACTGAG 420
 DB 402 AAGTTGAAACCCCGGAAGATACCATGATCATGAAGAGAGGAGAGTCACTGAG 461
 QY 421 CACATCCAGCGCATGACCGGAGCTTAAGCGGACATTAAGAGACATTAATTTTCAAG 480
 DB 462 CACATCCAGCGCATGACCGGAGCTTAAGCGGACATTAAGAGACATTAATTTTCAAG 521
 QY 481 GATGATTCGGAACCAAGCAGCGGAGCACTCACTCTCTGAGCATATGAGAGAT 540
 DB 522 GATGATTCGGAACCAAGCAGCGGAGCACTCACTCTCTGAGCATATGAGAGAT 581
 QY 541 AACCCGAGGTGGGCTACTGACGAGACCTGAGCCATGCGCTTGTCTCTCTAT 600
 DB 582 AACCCGAGGTGGGCTACTGACGAGACCTGAGCCATGCGCTTGTCTCTCTAT 641
 QY 601 CTTCTGAGAGAGATGCAATTCTGGGCACTGGTCACTGTGCTGGCACTGAGAGGCACTCC 660
 DB 642 CTTCTGAGAGAGATGCAATTCTGGGCACTGGTCACTGTGCTGGCACTGAGAGGCACTCC 701
 QY 661 CTGACGAGATTTCAAGCCCAAAATGGCGGAGCCGTCAGAGGAGCTCAAGACCAAGAG 720

Db 702 CTGACGGAGATTTCACAGCCCAATGCGGGACCTGCCAGGGGCTCAAGACCAACAGAG 761
Qy 721 CATGTGTAGCAAGCTGACAAACCCAGAGCAATGGGGATGAGACAAAGAAATCTATGT 780
Db 762 CATGTGTAGCAAGCTGACAAACCCAGAGCAATGGGGATGAGACAAAGAAATCTATGT 821
Qy 781 GGGCAGTGTCCCGTGTAGGCTGCTCATCGGATATTTGATTGAAGGGAATCTCTCGGG 840
Db 822 GGGCAGTGTCCCGTGTAGGCTGCTCATCGGATATTTGATTGAAGGGAATCTCTCGGG 881
Qy 841 CTGACCTGCGCTGTGGAGCTGTATCTGTGTAGAGAGCGCAACAGGGCTGTATGCCATA 900
Db 882 CTGACCTGCGCTGTGGAGCTGTATCTGTGTAGAGAGCGCAACAGGGCTGTATGCCATA 941
Qy 901 ACAAGATGCGCTTTAAGTTCAGCAGAGGCTCTCAAGAGAGCTCAGGTGAGGCCG 960
Db 942 ACAAGATGCGCTTTAAGTTCAGCAGAGGCTCTCAAGAGAGCTCAGGTGAGGCCG 1001
Qy 961 TGGGCACTTTTTCACACCGGTTCTGTTGATCTGAGGCAAGGAGATGAGACACTGTGCTC 1020
Db 1002 TGGGCACTTTTTCACACCGGTTCTGTTGATCTGAGGCAAGGAGATGAGACACTGTGCTC 1061
Qy 1021 AAGCATTTAAGGCTCTTATGAGAACTTAAAGAAAGAGAGGAGACCTGCCACCCCA 1080
Db 1062 AAGCATTTAAGGCTCTTATGAGAACTTAAAGAAAGAGAGGAGACCTGCCACCCCA 1121
Qy 1081 GCCAAACCCAGAGAGGCTGTGGCACTGAGGCTGTGGCCGCTTCAAGTGGCGGGAG 1140
Db 1122 GCCAAACCCAGAGAGGCTGTGGCACTGAGGCTGTGGCCGCTTCAAGTGGCGGGAG 1181
Qy 1141 ACCCTCTGCAAGGGGGACAGGCAAGGCCCTCTCAAGGCCCAAGCCCGGCTCCCGGGCC 1200
Db 1182 ACCCTCTGCAAGGGGGACAGGCAAGGCCCTCTCAAGGCCCAAGCCCGGCTCCCGGGCC 1241
Qy 1201 ATTGTGTCAGCTTCCCGCCCAAGGCACTCTGTTCTTCCACACCTCTGCTGAGGGGCT 1260
Db 1242 ATTGTGTCAGCTTCCCGCCCAAGGCACTCTGTTCTTCCACACCTCTGCTGAGGGGCT 1301
Qy 1261 GTCCGGGAGAGACCTTACCTGCTGGGCACTCAGGGGTGTGGCCAGCCCGGCTGTGGCTCAG 1320
Db 1302 GTCCGGGAGAGACCTTACCTGCTGGGCACTCAGGGGTGTGGCCAGCCCGGCTGTGGCTCAG 1361
Qy 1321 GGAAGACCTCAGGGGTTCTGTGAGATTCTGTGAGTGTCAATCCATGCCCGCTTCCCAAG 1380
Db 1362 GGAAGACCTCAGGGGTTCTGTGAGATTCTGTGAGTGTCAATCCATGCCCGCTTCCCAAG 1421
Qy 1381 GACCTGAGCTGAGAGGCGCTTGTGCTCCCATTTATGATTTTCAAGACAGAGCTGTGGGTC 1440
Db 1422 GACCTGAGCTGAGAGGCGCTTGTGCTCCCATTTATGATTTTCAAGACAGAGCTGTGGGTC 1481
Qy 1441 CGTGCATATTCCTCCAGAGAGACAGCTGGCCCTGTGGCGAGCTGAAACCTCTGGGGAG 1500
Db 1482 CGTGCATATTCCTCCAGAGAGACAGCTGGCCCTGTGGCGAGCTGAAACCTCTGGGGAG 1541
Qy 1501 CGGCTGAGATCGGCTTTCGCTGCAACCAAGCACTGATTTCCGACCAAGGCAACCCCTTCA 1560
Db 1542 CGGCTGAGATCGGCTTTCGCTGCAACCAAGCACTGATTTCCGACCAAGGCAACCCCTTCA 1601
Qy 1561 GCTAGGAGCAAGACGCGGTGTCTCCCACTCAGGGGCTTGTGCGGCTTCACTTGTG 1620
Db 1602 GCTAGGAGCAAGACGCGGTGTCTCCCACTCAGGGGCTTGTGCGGCTTCACTTGTG 1661
Qy 1621 GAAAGTCTCAGTTCCTCCAGAGCTTCTAAGAGATCTGGGCGAGGGCTCAATGCTGGAT 1680
Db 1662 GAAAGTCTCAGTTCCTCCAGAGCTTCTAAGAGATCTGGGCGAGGGCTCAATGCTGGAT 1721
Qy 1681 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTGCCTCTGTTTTATTTTGTGTTAA 1740
Db 1722 AATTTCCCTAGGCTTAAACAACCAAGCAAG-TTGCGATCTCTGTTTTATTTTGTGTTAA 1780
Qy 1741 CTTATGAAAATGTATTTAAGAAAGTGCAGCTCAGAGAGATTCAGAGATGGAACACACC 1800
Db 1781 CTTATGAAAATGTATTTAAGAAAGTGCAGCTCAGAGAGATTCAGAGATGGAACACACC 1840

Qy 1801 AGACCCAGATCAAAAGCCAAACCATGCCAGCCCTCCAGAGACCCCAAGCCCAAGAC 1860
Db 1841 AGACCCAGATCAAAAGCCAAACCATGCCAGCCCTCCAGAGACCCCAAGCCCAAGAC 1900
Qy 1861 CATGCTTCTGAATTTCTGACGACACCGTGAGCTGCTGTTGACTTTAACTTATGGAAG 1920
Db 1901 CATGCTTCTGAATTTCTGACGACACCGTGAGCTGCTGTTGACTTTAACTTATGGAAG 1960
Qy 1921 ATAACTAACCTTCACGTTTGAATAATGTTTC 1953
Db 1961 ATAACTAACCTTCACGTTTGAATAATGAGTC 1993

RESULT 9
AD37380
ID AD37380 standard, DNA; 2304 BP.
XX
AC AD37380;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 213.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Neotropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WC0203048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002MO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
XX
DR P-PSDB; AD37381.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancer,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 213; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (AD37168-AD37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancer, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 2304 BP; 569 A; 610 C; 636 G; 489 T; 0 U; 0 Other;

Query Match 96.8%; Score 1902; DB 10; Length 2304;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1937; Conservative 0; Mismatches 5; Indels 22; Gaps 1;

Qy 1 ATGAGCTGTGTAGAGTTCGCGGCACTTGTGGGCAAGAGGAGGACATCATTTATG 60
Db 362 ATGAGCTGTGTAGAGTTCGCGGCACTTGTGGGCAAGAGGAGGACATCATTTATG 421

QY 61 AAATAGAAAGGAGACCGAGCTGGGCTGCGAGAGACAAAGGGGCTTAAGCTTTTCCA 120
 DB 422 AAATAGAAAGGAGACCGAGCTGGGCTGCGAGAGACAAAGGGGCTTAAGCTTTTCCA 481
 QY 121 AGCTACAAACAACGTCGATCATTTTGGGATTTGTATCATGAGACGAGTCCCTCTG 180
 DB 482 AGCTACAAACAACGTCGATCATTTTGGGATTTGTATCATGAGACGAGTCCCTCTG 541
 QY 181 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAAGCCGAAAGCAAGTGGTGAAT 240
 DB 542 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAAGCCGAAAGCAAGTGGTGAAT 601
 QY 241 ATGCTGGAGACTGGAGAAATACAAAGCAGCAAGGCTCATAGATGAGGCTCAAG 300
 DB 602 ATGCTGGAGACTGGAGAAATACAAAGCAGCAAGGCTCATAGATGAGGCTCAAG 661
 QY 301 GGAATGCCATGAACATCCGGGGCCGATGTGTCAGTCTCTCTGAACATTTGAGAAATG 360
 DB 662 GGAATGCCATGAACATCCGGGGCCGATGTGTCAGTCTCTCTGAACATTTGAGAAATG 721
 QY 361 AAGTTGAAAAACCCCGGAAGATTCAGATCATGAGAGAGAGGCAAGGTCTTGAAG 420
 DB 722 AAGTTGAAAAACCCCGGAAGATTCAGATCATGAGAGAGAGGCAAGGTCTTGAAG 781
 QY 421 CACATCAGGCGATGACCGGGGACGTAAAGCGGGGCAATTAAAGGAAGATATATCTTCAGG 480
 DB 782 CACATCAGGCGATGACCGGGGACGTAAAGCGGGGCAATTAAAGGAAGATATATCTTCAGG 841
 QY 481 GATCGATACGGAACCAAGCAGCGGGGAACTACTCAGATCCTCTCTGATATGAGAGAT 540
 DB 842 GATCGATACGGAACCAAGCAGCGGGGAACTACTCAGATCCTCTCTGATATGAGAGAT 901
 QY 541 AACCGGAGGTGGGCTACTGCAAGGACCTGAGCCATGCGCGCTTGTCTCTCTAT 600
 DB 902 AACCGGAGGTGGGCTACTGCAAGGACCTGAGCCATGCGCGCTTGTCTCTCTAT 961
 QY 601 CTTCCTGAGAGGATGATCTTGGGGCACTGGGACGTCGTCGTCGAGTGAAGGACCTCC 660
 DB 962 CTTCCTGAGAGGATGATCTTGGGGCACTGGGACGTCGTCGTCGAGTGAAGGACCTCC 1021
 QY 661 CTGCAAGGATTTCAACGCCCCAAATGGCGGGACCTGTCAGAGGGCTCCAAAGCAACAGAG 720
 DB 1022 CTGCAAGGATTTCAACGCCCCAAATGGCGGGACCTGTCAGAGGGCTCCAAAGCAACAGAG 1081
 QY 721 CATGTGTATGCCACTGTCACAAACCCAAATGCGGCGCATCAGAGCAAGAAATCTATGT 780
 DB 1082 CATGTGTATGCCACTGTCACAAACCCAAATGCGGCGCATCAGAGCAAGAAATCTATGT 1141
 QY 781 GGGCAGTGTCCCGCTTAGGCTGCTCCTCAATCCGGATTTTGAATGACGGGATCTCTCGGG 840
 DB 1142 GGGCAGTGTCCCGCTTAGGCTGCTCCTCAATCCGGATTTTGAATGACGGGATCTCTCGGG 1201
 QY 841 CTCAACCTGCGCTGTGGGACGTGTATCTGTATGAAAGGCAAGCGCTTGAATGCGGATA 900
 DB 1202 CTCAACCTGCGCTGTGGGACGTGTATCTGTATGAAAGGCAAGCGCTTGAATGCGGATA 1261
 QY 901 ACAAGATTCGCTTTAAGGTTCAAGAAAGCGCTTCAGAAAGCGTCAAGGTGTGGCCCG 960
 DB 1262 ACAAGATTCGCTTTAAGGTTCAAGAAAGCGCTTCAGAAAGCGTCAAGGTGTGGCCCG 1321
 QY 961 TGGGCACTTTTTCGAAACCGGTTCTGTTGATACCTGGGCGAAGGATGAGCACTGTGCTC 1020
 DB 1322 TGGGCACTTTTTCGAAACCGGTTCTGTTGATACCTGGGCGAAGGATGAGCACTGTGCTC 1381
 QY 1021 AAGCATCTTAAGGCGCTCTATGAAGAAATCAACAAGAAACAGAGGGGACTGTGCAACCCCA 1080
 DB 1382 AAGCATCTTAAGGCGCTCTATGAAGAAATCAACAAGAAACAGAGGGGACTGTGCAACCCCA 1441
 QY 1081 GCCAAACCCGAGCAAGAGGTGTGTGCGCATCAGGCGCTGTGCGGCTTCACTGTGCGGGAAG 1140
 DB 1442 GCCAAACCCGAGCAAGAGGTGTGTGCGCATCAGGCGCTGTGCGGCTTCACTGTGCGGGAAG 1501

QY 1141 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCCAGAGGCCACAGCCCGGTTCCCGCGGCCC 1200
 DB 1502 ACCCTTTCAGAGGGGAGCAGGAGGCGCCCTCCA-----GGCCC 1539
 QY 1201 ATTGGTCAGCTTCCCGCAGCGGACCTGTTCTTCACAACCTGTCTGTGGGGCT 1260
 DB 1540 ATTGGTCAGCTTCCCGCAGCGGACCTGTTCTTCACAACCTGTCTGTGGGGCT 1599
 QY 1261 GTCCGGGAAGACACTACCTGTGGGACCTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
 DB 1600 GTCCGGGAAGACACTACCTGTGGGACCTCAGGGTGTGCCAGCCCGGCTGCTCAG 1659
 QY 1321 GGAGGACCTCAGGGTTCCTGAGATTCCTGCAAGTGAATCCGATGCGCGGCTCCAAAG 1380
 DB 1660 GGAGGACCTCAGGGTTCCTGAGATTCCTGCAAGTGAATCCGATGCGCGGCTCCAAAG 1719
 QY 1381 GACCTGACGTAGAGGCGCTTGTGTCGCAATTATGATTTTCAACAGAGCTGCTGGTTC 1440
 DB 1720 GACCTGACGTAGAGGCGCTTGTGTCGCAATTATGATTTTCAACAGAGCTGCTGGTTC 1779
 QY 1441 CGTCCATATCCAGAGGAGCAGAGTGGCCCCCTGTGGCAGGCTGGAACACCTGCGGAG 1500
 DB 1780 CGTCCATATCCAGAGGAGCAGAGTGGCCCCCTGTGGCAGGCTGGAACACCTGCGGAG 1839
 QY 1501 CGGGTGAAGATCGGCTTTGCTGCAACCCAGCACTGAATCCGACAGGGGCAACCCCTTCAG 1560
 DB 1840 CGGGTGAAGATCGGCTTTGCTGCAACCCAGCACTGAATCCGACAGGGGCAACCCCTTCAG 1899
 QY 1561 GCTAAGGAGCAAGCGGCTGTGCTCCACCTCAGGAGCTTGTGCTGCGGCTCCACTTG 1620
 DB 1900 GCTAAGGAGCAAGCGGCTGTGCTCCACCTCAGGAGCTTGTGCTGCGGCTCCACTTG 1959
 QY 1621 GAAAGTTCTCAGTTCCTTCAGGCTTTCAGAGCATCTGGGCCAGGGCTCATGCTGAT 1680
 DB 1960 GAAAGTTCTCAGTTCCTTCAGGCTTTCAGAGCATCTGGGCCAGGGCTCATGCTGAT 2019
 QY 1681 AATTTCCCTAAGGCTTAAACAACCCAGCAAGCTTGGCTGCTGTTTATTTTGGTTAA 1740
 DB 2020 AATTTCCCTAAGGCTTAAACAACCCAGCAAGCTTGGCTGCTGTTTATTTTGGTTAA 2079
 QY 1741 CTTATGAATAATATTTAAGAAAGATGAGCTCGAGAGAGATTCAGAGTGAACAACACC 1800
 DB 2080 CTTATGAATAATATTTAAGAAAGATGAGCTCGAGAGAGATTCAGAGTGAACAACACC 2139
 QY 1801 AGACCCAGATCAAAAGCCAAATGCCAGCCCTCTCCAGACACCCCAAGCCCAAGAC 1860
 DB 2140 AGACCCAGATCAAAAGCCAAATGCCAGCCCTCTCCAGACACCCCAAGCCCAAGAC 2199
 QY 1861 CATCGTTCTGAATTTCTGACGACACCGTGAGCTGCTTGTATCTTAACTCATGGAAG 1920
 DB 2200 CATCGTTCTGAATTTCTGACGACACCGTGAGCTGCTTGTATCTTAACTCATGGAAG 2259
 QY 1921 ATTAACCTTCAAGCTTTTGAATAATGTTTCTGTGAATG 1964
 DB 2260 ATTAACCTTCAAGCTTTTGAATAATGTTTCTGTGAATG 2303

RESULT 10
 ADC37378
 ID ADC37378 standard; DNA; 2304 BP.
 XX
 AC ADC37378;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 211.
 XX
 KW Nuclear factor kappa B, NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
 KW Immunomodulator; Cytostatic; Anticancer; Osteoporosis; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiac; Gene therapy; human; gene; ds.

OS Homo sapiens.
 XX
 PN WO2003048202-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-JP012644.
 XX
 PR 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 XX (ASAH) ASAH KASEI KK.
 PA
 PI Matsuda A, Muramatsu S;
 XX
 DR WPI; 2003-505282/47.
 DR P-PSDB; ADC37379.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX
 PS Claim 4; SEQ ID NO 211; 938pp; English.

CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischemic disorders.

XX Sequence 2304 BP; 570 A; 609 C; 636 G; 489 T; 0 U; 0 Other;

Query Match 96.8%; Score 1900.4; DB 10; Length 2304;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1936; Conservative 0; Mismatches 6; Indels 22; Gaps 1;

QY 1 ATGAGCGTGTAGAGTCCGCGGCACTTGTGGGACAAAGACGAGACATCATATG 60
 DB 362 ATGAGCGTGTAGAGTCCGCGGCACTTGTGGGACAAAGACGAGACATCATATG 421
 QY 61 AAATACGAAAAGGACACCGAGCTGGGCTCCAGAGGACAAAGGGGCTTAAGCCTTTTGA 120
 DB 422 AAATACGAAAAGGACACCGAGCTGGGCTCCAGAGGACAAAGGGGCTTAAGCCTTTTGA 481
 QY 121 AGCTACAAACAACAGTGCATCTTTGGGATTGTACATGAGACGAGCTCCTCTG 180
 DB 482 AGCTACAAACAACAGTGCATCTTTGGGATTGTACATGAGACGAGCTCCTCTG 541
 QY 181 ACTGCGGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAAGACAAAGTGGTGGAT 240
 DB 542 ACTGCGGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAAGACAAAGTGGTGGAT 601
 QY 241 ATCTGCGGAACTCGGAGAAATACAAAGACGAGAAAGCTCATAGATCGAGGTAAAG 300
 DB 602 ATCTGCGGAACTCGGAGAAATACAAAGACGAGAAAGCTCATAGATCGAGGTAAAG 661
 QY 301 GGAATGCGCATGAAACATCCGGGGCCGATGTGTCACTCTCTGAACATTTGAGGAATG 360
 DB 662 GGAATGCGCATGAAACATCCGGGGCCGATGTGTCACTCTCTGAACATTTGAGGAATG 721
 QY 361 AAGTTGAAAAACCCCGGAAGATACAGATCATGAGGAGAAAGGGCAAGGTCATCTGAG 420
 DB 722 AAGTTGAAAAACCCCGGAAGATACAGATCATGAGGAGAAAGGGCAAGGTCATCTGAG 781
 QY 421 CAGATCCAGGCACTCGACCGGAGACGTAAAGCGGACATTTAAGAAAGCATATTTCTTAG 480
 DB 782 CAGATCCAGGCACTCGACCGGAGACGTAAAGCGGACATTTAAGAAAGCATATTTCTTAG 841
 QY 481 GATGATACGGAACCAAGACGCGGAACCTACTCACAATCTCTGCGATATGAGAGTAT 540

DB 842 GATGATACGGAACCAAGACGCGGAACCTACTCACAATCTCTGCGATATGAGAGTAT 901
 QY 541 AACCCGGAAGTGGGCTACGTCAGGGGACCTGAGGACATAGCGGCTTTGCTCTCTAT 600
 DB 902 AACCCGGAAGTGGGCTACGTCAGGGGACCTGAGGACATAGCGGCTTTGCTCTCTAT 961
 QY 601 CTTCCTGAGGAGGATGATCTTGGGCACTGTGTCACTGTGCGCATGAGAGGACATCC 660
 DB 962 CTTCCTGAGGAGGATGATCTTGGGCACTGTGTCACTGTGCGCATGAGAGGACATCC 1021
 QY 661 CTTCAGGAGATTTCAACGCCCAATGCGGAGACCTGCAGGGGCTCCAAAGCAACGAGAG 720
 DB 1022 CTTCAGGAGATTTCAACGCCCAATGCGGAGACCTGCAGGGGCTCCAAAGCAACGAGAG 1081
 QY 721 CATGTGTAGCCACCTGACCAACCCAGACATGAGGACATGAGCAAAAGATCTATGT 780
 DB 1082 CATGTGTAGCCACCTGACCAACCCAGACATGAGGACATGAGCAAAAGATCTATGT 1141
 QY 781 GGGCAGTGTCCCGTTAAGGCTGCTCATCCGGATATGATTTGACGGGATCTCTCGGG 840
 DB 1142 GGGCAGTGTCCCGTTAAGGCTGCTCATCCGGATATGATTTGACGGGATCTCTCGGG 1201
 QY 841 CTGACCTTGCGCTGTGGGACGTGATCTGTGAAAGCGGAACAGGCGTTGATGCCGATA 900
 DB 1202 CTGACCTTGCGCTGTGGGACGTGATCTGTGAAAGCGGAACAGGCGTTGATGCCGATA 1261
 QY 901 ACAAGATGCGCTTTAAGGCTTACGAGGAGCGCTCACAAAGCCTGAGGTGTGCGCG 960
 DB 1262 ACAAGATGCGCTTTAAGGCTTACGAGGAGCGCTCACAAAGCCTGAGGTGTGCGCG 1321
 QY 961 TGGGCAGTGTTTTGAACAGCGGTTCGTTGATACCTGGGACAGGATGAGACATGTGCTC 1020
 DB 1322 TGGGCAGTGTTTTGAACAGCGGTTCGTTGATACCTGGGACAGGATGAGACATGTGCTC 1381
 QY 1021 AAGCATCTTAAGGCGCTCTATGAAAGAACTTAACAAGAACGAGGGGACCTGCCACCCCA 1080
 DB 1382 AAGCATCTTAAGGCGCTCTATGAAAGAACTTAACAAGAACGAGGGGACCTGCCACCCCA 1441
 QY 1081 GCCAAACCCGAGAACAGGTCGTGGGATCCAGGCGTGTCCGCTTACGTTGGCGGGAAG 1140
 DB 1442 GCCAAACCCGAGAACAGGTCGTGGGATCCAGGCGTGTCCGCTTACGTTGGCGGGAAG 1501
 QY 1141 ACCCTCTGCAAGGGGACAGGCGGCGCTTCAGGCGCCACAGCGCGTTCGCGGCGCC 1200
 DB 1502 ACCCTCTGCAAGGGGACAGGCGGCGCTTCAGGCGCGCCACAGCGCGTTCGCGGCGCC 1539
 QY 1201 ATTGTGACGCTTCCCGGCCACGCGGACCTGCTTTTCCACACCTGTCTGTGGGAGCT 1260
 DB 1540 ATTGTGACGCTTCCCGGCCACGCGGACCTGCTTTTCCACACCTGTCTGTGGGAGCT 1599
 QY 1261 GTCCGGGAAGACACTAACCTGTGGGACATCAGGGGTGTCGCCAGCCCGGCGCTGGCTCAG 1320
 DB 1600 GTCCGGGAAGACACTAACCTGTGGGACATCAGGGGTGTCGCCAGCCCGGCGCTGGCTCAG 1659
 QY 1321 GAGAGACCTCAAGGCTTCTGAGAAATCTGTGAGTGAACCTCAATGCCCCGCTCCCAAG 1380
 DB 1660 GAGAGACCTCAAGGCTTCTGAGAAATCTGTGAGTGAACCTCAATGCCCCGCTCCCAAG 1719
 QY 1381 GACCTGACGTAAGGAGGCGCTTGGTCCGCAATTAATGATTTCAACAAGCTGTGGGCTC 1440
 DB 1720 GACCTGACGTAAGGAGGCGCTTGGTCCGCAATTAATGATTTCAACAAGCTGTGGGCTC 1779
 QY 1441 CGTGCCATATCCAGAGAGACAGCTGTGCGCCCTGTGTCAGGCTGAACAACCCCTGGGAG 1500
 DB 1780 CGTGCCATATCCAGAGAGACAGCTGTGCGCCCTGTGTCAGGCTGAACAACCCCTGGGAG 1839
 QY 1501 CGGGTGAAGATCGGCTTTGCTGCAACCAAGCATGATTCGACAGGAGACCCCTTCAGA 1560
 DB 1840 CGGGTGAAGATCGGCTTTGCTGCAACCAAGCATGATTCGACAGGAGACCCCTTCAGA 1899
 QY 1561 GCTAAGGAGCAAGCAGCGTGTGCTCCACCTCAAGGCGCTTGTGCGGCTTCACCTTG 1620

DB 1900 GCTAGGAGCAACAGCCGCTGCTGCCACCTCAGGCGCTTGCTGCGGCTCCACTTG 1959
QY 1621 GAAAGTTCAGTTCCTCCAGGCTTCTGAAGCATCTGGGCCAGGCTCAGTGGCTGAT 1680
DB 1960 GAAAGTTCAGTTCCTCCAGGCTTCTGAAGCATCTGGGCCAGGCTCAGTGGCTGAT 2019
QY 1681 AATTTCCCTAGGCTTAAACCAACCAAGCTTCGGGCTCTCTTTATTTTGGTTAAA 1740
DB 2020 AATTTCCCTAGGCTTAAACCAACCAAGCTTCGGGCTCTCTTTATTTTGGTTAAA 2079
QY 1741 CTTATGAAAAATGATTTAAGAAAGAGTGCAGCTCGAGAGATTCAGAGATGAAACACACC 1800
DB 2080 CTTATGAAAAATGATTTAAGAAAGAGTGCAGCTCGAGAGATTCAGAGATGAAACACACC 2139
QY 1801 AGACCCAGATGCAAGAGCAACCATGCCCCCTCCAGAGACCCCAAGCCCAAGCAGAC 1860
DB 2140 AGACCCAGATGCAAGAGCAACCATGCCCCCTCCAGAGACCCCAAGCCCAAGCAGAC 2199
QY 1861 CATCGTTCTGAATTCGACGACACCGTGAAGCTGCTTTGTAATTTAACTCATGGAAG 1920
DB 2200 CATCGTTCTGAATTCGACGACACCGTGAAGCTGCTTTGTAATTTAACTCATGGAAG 2259
QY 1921 AATACTACTTCACTGTTTGAATAATGTTTCTGTTGAATG 1964
DB 2260 AATACTACTTCACTGTTTGAATAATGTTTCTGTTGAATG 2303

RESULT 11
AB211810
ID AB211810 standard; cDNA; 2084 BP.
XX AB211810;
AC
XX
XX 20-JAN-2003 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 692.
DE
XX
XX Human: genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antichronic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200270539-A2.
PN
XX
XX 12-SEP-2002.
PD
XX
XX 05-MAR-2002; 2002WO-US005095.
PF
XX
XX 05-MAR-2001; 2001US-00799451.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YF, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao Q, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehman T, Wang J, Wang D, Drmanac RT;
XX
XX MPI: 2002-759812/82.
DR P-PSDB; ABP69593.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
XX Claim 1; SEQ ID NO 692; 1012bp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences (AB21119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burn, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;
SQ
Query Match 93.5%; Score 1836.4; DB 6; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

QY 1 ATGACGTGTAGAGGTGCGGGCAATTGTTGGGCAAGAGCGAGAGCATCATATG 60
DB 163 ATGACGTGTAGAGGTGCGGGCAATTGTTGGGCAAGAGCGAGAGCATCATATG 222
QY 61 AATATGAAAAAGGACACCGAGTGGCTGCCAGAGGAGCAAGGGGCTTAAGCTTTTGA 120
DB 223 AATATGAAAAAGGACACCGAGTGGCTGCCAGAGGAGCAAGGGGCTTAAGCTTTTGA 282
QY 121 AGCTACAAACAAAGTGCATTTGGGGATTTGATCATAGAGCGAGGCTGCTCTG 180
DB 283 AGCTACAAACAAAGTGCATTTGGGGATTTGATCATAGAGCGAGGCTGCTCTG 342
QY 181 ACTGCGCGGAGCGAGCAAGCAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGTAT 240
DB 343 ACTGCGCGGAGCGAGCAAGCAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGTAT 402
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAGACAGAAAGCTCATATGAGCTGATCAG 300
DB 403 ATGCTGGGAGACTGGGAGAAATACAAAGACAGAAAGCTCATATGAGCTGATCAG 462
QY 301 GGAAATGCCATTAACATCCGGGGCCGAGTGTGTGATCTCTCTGAACATTAGAGAAATG 360
DB 463 GGAAATGCCATTAACATCCGGGGCCGAGTGTGTGATCTCTCTGAACATTAGAGAAATG 522
QY 361 AAGTTGAAAAAACCAGGAAGATACAGATCATGAAGAGAGGCAAGATCATCTGAG 420
DB 523 AAGTTGAAAAAACCAGGAAGATACAGATCATGAAGAGAGGCAAGATCATCTGAG 529
QY 421 CACATCCAGCGCATGACCGGGAGCTTAAGCGGAGCATTAAGAGCATATATTCTTCAAG 480
DB 530 -ACATCCAGCGCATGACCGGGAGCTTAAGCGGAGCATTAAGAGCATATATTCTTCAAG 588
QY 481 GATGATACGGAAACCAAGAGCGGAGACTACTCCATCTCTCTGCAATATAGAGATAT 540
DB 589 GATGATACGGAAACCAAGAGCGGAGACTACTCCATCTCTCTGCAATATAGAGATAT 648
QY 541 AACCCGAGTGGGTACTGACGAGGACCTGAGCCATATGCGCGCTTGTCTCTCTAT 600
DB 649 AACCCGAGTGGGTACTGACGAGGACCTGAGCCATATGCGCGCTTGTCTCTCTAT 708
QY 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGACGTGCTGCGCAATGAGGCACTCC 660
DB 709 CTTCTGAGAGAGATGATTTCTGGGCACTGTGACGTGCTGCGCAATGAGGCACTCC 768
QY 661 CCGAGGAGTTTCAAGCCAAATGCGGAGACGTCGAGGGCTCCAGAGCAAGAGAG 720
DB 769 CCGAGGAGTTTCAAGCCAAATGCGGAGACGTCGAGGGCTCCAGAGCAAGAGAG 828
QY 721 CATGTGTAGCCAGCTCAACCAAGACCATGAGGAGCATCAGAGCAAGAAAGATCATGT 780
DB 829 CATGTGTAGCCAGCTCAACCAAGACCATGAGGAGCATCAGAGCAAGAAAGATCATGT 888


```

Qy 781 GGGCAGTGTCCCGTTAGGCTGCCTCATCCGATATTGATTTGACGGGATCTCTCTGGG 840
    |||
Db 889 GGGCAGTGTCCCGTTAGGCTGCCTCATCCGATATTGATTTGACGGGATCTCTCTGGG 948
Qy 841 CTGACCTCTGCGCTCTGTGGAGCTGTATCTGGTAGAAGCCGAAACAGCGTTGATGCCGATA 900
    |||
Db 949 CTGACCTCTGCGCTCTGTGGAGCTGTATCTGGTAGAAGCCGAAACAGCGTTGATGCCGATA 1008
Qy 901 ACAAGATCCGCTTTAAGTTCAGCAGAAAGCGCTCAGAAAGCTCAGGTGGGCGCG 960
    |||
Db 1009 ACAAGATCCGCTTTAAGTTCAGCAGAAAGCGCTCAGAAAGCTCAGGTGGGCGCG 1068
Qy 961 TGGGCACTTTTTCGCAACCGGTTCTGTGATATCTGGGCCAGGAGTAGAGCACTGTGCTC 1020
    |||
Db 1069 TGGGCACTTTTTCGCAACCGGTTCTGTGATATCTGGGCCAGGAGTAGAGCACTGTGCTC 1128
Qy 1021 AAGCATCTTAGGGCTCTATAGAAAGAACTAAAGAAAGAGGGGAACTTGCCACCCCA 1080
    |||
Db 1129 AAGCATCTTAGGGCTCTATAGAAAGAACTAAAGAAAGAGGGGAACTTGCCACCCCA 1188
Qy 1081 GCCAAACCCGAGCAAGGCTGTCTGGCATCCAGGCTGTGCCGCTTCACTGTGGGGAG 1140
    |||
Db 1189 GCCAAACCCGAGCAAGGCTGTCTGGCATCCAGGCTGTGCCGCTTCACTGTGGGGAG 1248
Qy 1141 ACCCTTGCAAGGGGAGCAAGGAGGCCCTTCAGAGCCCAACAGGCCGATTCGCGCGGCC 1200
    |||
Db 1249 ACCCTTGCAAGGGGAGCAAGGAGGCCCTTCAGAGCCCAACAGGCCGATTCGCGCGGCC 1308
Qy 1201 ATTTGTCAGCTTCCCGCCCAACCGGCACTCTGTTCTTCCACACCTCTGCTGTGGGCT 1260
    |||
Db 1309 ATTTGTCAGCTTCCCGCCCAACCGGCACTCTGTTCTTCCACACCTCTGCTGTGGGCT 1368
Qy 1261 GTCCGGGAAGACACTACCTCTGTGGGACATCCAGGCTGTGCCAGCCGCTGGGCTCAG 1320
    |||
Db 1369 GTCCGGGAAGACACTACCTCTGTGGGACATCCAGGCTGTGCCAGCCGCTGGGCTCAG 1428
Qy 1321 GGAGGACCTCAGGCTTCTGTGAGATCTCTGCAATGGAATCTCATGCCGCTTCCCAACG 1380
    |||
Db 1429 GGAGGACCTCAGGCTTCTGTGAGATCTCTGCAATGGAATCTCATGCCGCTTCCCAACG 1488
Qy 1381 GACCTGAGACGTAGAGGCGCTTGTGCTCCGCAATTATGATTTGACAGAGCTGTGGGCTC 1440
    |||
Db 1489 GACCTGAGACGTAGAGGCGCTTGTGCTCCGCAATTATGATTTGACAGAGCTGTGGGCTC 1548
Qy 1441 CGTGCCTATTCGCGAGAGACCAACTGGCCCCCTGTGCGAGGCTGAAACCCCTGCGAG 1500
    |||
Db 1549 CGTGCCTATTCGCGAGAGACCAACTGGCCCCCTGTGCGAGGCTGAAACCCCTGCGAG 1608
Qy 1501 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCGCAACGAGGCAACCCCTTGACA 1560
    |||
Db 1609 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCGCAACGAGGCAACCCCTTGACA 1668
Qy 1561 GCTAGGAGCAAGACGCGTGTGCTCCGACCTCAGGGCGCTTGCTGTGCGGCTTCCACTTG 1620
    |||
Db 1669 GCTAGGAGCAAGACGCGTGTGCTCCGACCTCAGGGCGCTTGCTGTGCGGCTTCCACTTG 1728
Qy 1621 GAAAGTCTCAGTTCCTCCGAGGCTTCTAAGAAAGATCTGGGCCAGGGCTCATGTGCTGAT 1680
    |||
Db 1729 GAAAGTCTCAGTTCCTCCGAGGCTTCTAAGAAAGATCTGGGCCAGGGCTCATGTGCTGAT 1788
Qy 1681 AATTTCCTAGGCTTAAACAACCAAGCAAGCTTGCGCTCGTATTTATTTTGGTTAAA 1740
    |||
Db 1789 AATTTCCTAGGCTTAAACAACCAAGCAAGCTTGCGCTCGTATTTATTTTGGTTAAA 1848
Qy 1741 CTATAGAAAATGTATTAAGAAAAGATGCAAGCTCAGAGAGATTTCAGAGATGAAACACACC 1800
    |||
Db 1849 CTATAGAAAATGTATTAAGAAAAGATGCAAGCTCAGAGAGATTTCAGAGATGAAACACACC 1908
Qy 1801 AGACCCCAAGATCAAAAGCCCAACATGCGCCCTTCCAGACACCCCAAGCCCAAGAC 1860
    |||
Db 1909 AGACCCCAAGATCAAAAGCCCAACATGCGCCCTTCCAGACACCCCAAGCCCAAGAC 1968

```

```

Qy 1861 CATCGTTCGAAATTCGACGACACCGTGAAGCTGCGCTTTGATCTTAAACTCATGGAAG 1920
    |||
Db 1969 CATCGTTCGAAATTCGACGACACCGTGAAGCTGCGCTTTGATCTTAAACTCATGGAAG 2028
Qy 1921 ATTAACCTTCACTTCACTTTTGAATAATGTTTCCCTTGAAATG 1964
    |||
Db 2029 ATTAACCACTTCATGTTTGAATAATGTTTCTGTGAATG 2072

RESULT 12
ADMA4328
ID ADMA4328 standard; cDNA; 2084 BP.
XX
AC ADMA4328;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #692.
XX
KM s6; gene; human; arginine-rich protein; cancer; inflammation;
KM genetic disorder.
XX
OS Homo sapiens.
XX
PN US2004053250-A1.
XX
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
XX 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US0005095.
PR 20-AUG-2002; 2002US-00225251.
XX
PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
PI Tang YT, Xue A, Drmanac RT;
XX
DR WPI; 2004-238579/22.
XX
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
XX Disclosure; SEQ ID NO 692; Sipp; English.
PS
XX
CC The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
XX
SQ Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Query Match 93.5%; Score 1836.4; DB 12; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

Qy 1 ATGAGCTGTGATAGAGTTCGGGCACTTGTGGGCAAGAGGAGAGCATATTATG 60
    |||
Db 163 ATGAGCTGTGATAGAGTTCGGGCACTTGTGGGCAAGAGGAGAGCATATTATG 222
Qy 61 AAATACGAAAAGGAGACACGAGCTGGGCTGCGAGAGAACAGGGGCTTAAGCTTTTGA 120
    |||
Db 223 AAATACGAAAAGGAGACACGAGCTGGGCTGCGAGAGAACAGGGGCTTAAGCTTTTGA 282
Qy 121 AGCTACAAACAACAGCTGATCTTTGGGGATTGTATATGAGACGAGAGCTGCTCTCTG 180

```


Db 283 ACCTACAAACAAGCTCATGTTGGGGAATTGTATCATGAGACGAGCTGCTCTCTG 342
Qy 181 ACTGCGCGGAGGCGGAAGCAATTCGCGCGAGATCAAGCCGAAAGCAATGGGTGAT 240
Db 343 ACTGCGCGGAGGCGGAAGCAATTCGCGCGAGATCAAGCCGAAAGCAATGGGTGAT 402
Qy 241 ATGCTGGGAGACTGGGAAATTAACAAGACGAGAAGCTCATATGAGGCTCAAG 300
Db 403 ATGCTGGGAGACTGGGAAATTAACAAGACGAGAAGCTCATATGAGGCTCAAG 462
Qy 301 GGAATGCCCATGACATCCGCGGCCCGGATGCTGCTCTCTGAAATGGAAGATG 360
Db 463 GGAATGCCCATGACATCCGCGGCCCGGATGCTGCTCTCTGAAATGGAAGATG 522
Qy 361 AGTTGAAAAAACCAGGAAGATACCAATCATGAAAGAGGCAAGAGTCATGAG 420
Db 523 AGTTGA----- 529
Qy 421 CACATCCAGCGCATGACCGGGAGCTAAGCGGACATTAAGGAAGCATATTTCTCAG 480
Db 530 -ACATCCAGCGCATGACCGGGAGCTAAGCGGACATTAAGGAAGCATATTTCTCAG 588
Qy 481 GATCGATACGGAACCAAGACGCGGAACTACTCAATCTCTGSCATATGAGAGTAT 540
Db 589 GATCGATACGGAACCAAGACGCGGAACTACTCAATCTCTGSCATATGAGAGTAT 648
Qy 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATGTCGCGCTTGTCTCTAT 600
Db 649 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATGTCGCGCTTGTCTCTAT 708
Qy 601 CTTCCTGAGGAGATGCAATTCGCGGACCTGTCAGCTCTGCGCATGAGAGCATCC 660
Db 709 CTTCCTGAGGAGATGCAATTCGCGGACCTGTCAGCTCTGCGCATGAGAGCATCC 768
Qy 661 CTGACGAGGATTCACAGCCCAATGCGGAGCGTCAAGGCTCCAAAGCAACAGAG 720
Db 769 CTGACGAGGATTCACAGCCCAATGCGGAGCGTCAAGGAGCTCCAAAGCAACAGAG 828
Qy 721 CATGTGTATGCAACGTCACAAACCAATGGGCAATCAAGCAAGAAATCTATGT 780
Db 829 CATGTGTATGCAACGTCACAAACCAATGGGCAATCAAGCAAGAAATCTATGT 888
Qy 781 GGGCAGTGTCCCGCTAGGCTGCTCATCCGATATTAATGACGGATCTCTCTCGG 840
Db 889 GGGCAGTGTCCCGCTAGGCTGCTCATCCGATATTAATGACGGATCTCTCTCGG 948
Qy 841 CTCAACCTGCGCTGTGAGACGTATCTGTAGAAAGCGAAACAGGCTTGAATGCGATA 900
Db 949 CTCAACCTGCGCTGTGAGACGTATCTGTAGAAAGCGAAACAGGCTTGAATGCGATA 1008
Qy 901 ACAAGAAATGCTTTAAGGTTGACAGAAAGCGCTTCAAGAGCGTCAAGGTGCGCCG 960
Db 1009 ACAAGAAATGCTTTAAGGTTGACAGAAAGCGCTTCAAGAGCGTCAAGGTGCGCCG 1068
Qy 961 TGGGACGTTTTTGAACCGGTTCTGTGATTAATCTGAGGCAAGGATGAGCACTGTGCTC 1020
Db 1069 TGGGACGTTTTTGAACCGGTTCTGTGATTAATCTGAGGCAAGGATGAGCACTGTGCTC 1128
Qy 1021 AAGCATCTTAAGGCTCTATGAAGAACTAACAAGAAAGCAGGAGGACCTGCAACCCCA 1080
Db 1129 AAGCATCTTAAGGCTCTATGAAGAACTAACAAGAAAGCAGGAGGACCTGCAACCCCA 1188
Qy 1081 GCCAAACCCGAGCAAGGCTGTGCGCATCCAGGCTGTGCGGCTTCACTGTGCGGAG 1140
Db 1189 GCCAAACCCGAGCAAGGCTGTGCGCATCCAGGCTGTGCGGCTTCACTGTGCGGAG 1248
Qy 1141 ACCCTCTGAAGGGGAGCAGGAGCCCTTCAAGGCGCCACAGCCGCTTCCCGGCGCC 1200
Db 1249 ACCCTCTGAAGGGGAGCAGGAGCCCTTCAAGGCGCCACAGCCGCTTCCCGGCGCC 1308
Qy 1201 ATTGGTCAGCTTCCCGGCAAGGCACTGCTTCTTCAACCTGCTGCTGAGGCT 1260

Db 1309 ATTGGTCAGCTTCCCGGCAAGGCACTGCTTCTTCCACACCTGCTGCTGAGGCT 1368
Qy 1261 GTCCGGAAAGACCTTACCTGTGGCACTCAGGGTGTGCCAGGCCGCTGCTCAG 1320
Db 1369 GTCCGGAAAGACCTTACCTGTGGCACTCAGGGTGTGCCAGGCCGCTGCTCAG 1428
Qy 1321 GGAAGGACCTCAGGGTTCCTGAGATTCCTGAGGTGAATTCATGCTGCGGCTCCAAAG 1380
Db 1429 GGAAGGACCTCAGGGTTCCTGAGATTCCTGAGGTGAATTCATGCTGCGGCTCCAAAG 1488
Qy 1381 GACTGAGCTGAGGGGCTTGTGCTCCGCAATTAATTAATTAATTAATTAATTAATTAAT 1440
Db 1489 GACTGAGCTGAGGGGCTTGTGCTCCGCAATTAATTAATTAATTAATTAATTAATTAAT 1548
Qy 1441 CGTGCATATCCAGAGGAGCAAGCTGCGCTGCTGAGGCTGAGCAACCTGCGAG 1500
Db 1549 CGTGCATATCCAGAGGAGCAAGCTGCGCTGCTGAGGCTGAGCAACCTGCGAG 1608
Qy 1501 CGGGTGAATGGGCTTGTGCTGCAACCGACATGATTCGACACAGGAGCAACCTGCTCAG 1560
Db 1609 CGGGTGAATGGGCTTGTGCTGCAACCGACATGATTCGACACAGGAGCAACCTGCTCAG 1668
Qy 1561 GCTAGGAGCAAGCAGCGGTGTGCTCCACCTCAGGAGCTTGTGCTGCGGCTCAGCTTG 1620
Db 1669 GCTAGGAGCAAGCAGCGGTGTGCTCCACCTCAGGAGCTTGTGCTGCGGCTCAGCTTG 1728
Qy 1621 GAAAGTCTCACTTCCCTCAGGCTTGAAGCATCTGGGCAAGGCTCATGCTGAT 1680
Db 1729 GAAAGTCTCACTTCCCTCAGGCTTGAAGCATCTGGGCAAGGCTCATGCTGAT 1788
Qy 1681 AATTTCCCTAGGCTTGAAGCAACCGACAGCTTCCGCTGCTTATTTTGTGTA 1740
Db 1789 AATTTCCCTAGGCTTGAAGCAACCGACAGCTTCCGCTGCTTATTTTGTGTA 1848
Qy 1741 CTTATGAATATTAAGAAAGATGAGCTCGAGAGATTCAGAGATGAAACACACC 1800
Db 1849 CTTATGAATATTAAGAAAGATGAGCTCGAGAGATTCAGAGATGAAACACACC 1908
Qy 1801 AGACCCAGATCAAAAGCAACATGCCAGCCCTCCAGACACCCCAAGCCCAAGC 1860
Db 1909 AGACCCAGATCAAAAGCAACATGCCAGCCCTCCAGACACCCCAAGCCCAAGC 1968
Qy 1861 CATGTTGCAATTCGACAGACCGTGAAGCTGCTTGTATCTTAATCATGAGAG 1920
Db 1969 CATGTTGCAATTCGACAGACCGTGAAGCTGCTTGTATCTTAATCATGAGAG 2028
Qy 1921 ATAACCTTCAAGCTTTTGAATATGTTCTCTGTTGAATG 1964
Db 2029 ATAACCTTCAATGTTTGAATATGTTCTCTGTTGAATG 2072

RESULT 13
ABZ11809
ID ABZ11809 standard; cDNA; 2072 BP.
XX
XX ABZ11809;
DT 20-JAN-2003 (first entry)
XX
XX Human polynucleotide SEQ ID NO 691.
DE
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cystostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antithrilitic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX
PN WO200270539-A2.

XX 12-SEP-2002.
 PD Best Local Similarity 92.3%; Score 1812.4; DB 6; Length 2072;
 XX 05-MAR-2002; 2002MO-US005095.
 PF Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
 XX 05-MAR-2001; 2001US-0079451.
 PR (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2002-759812/82.
 DR P-PSDB; ABP69592.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 XX
 XX Claim 1; SEQ ID NO 691; 1012bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (AB21119-
 CC AB212066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP68902-ABP68849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;
 Query Match 92.3%; Score 1812.4; DB 6; Length 2072;
 Best Local Similarity 96.3%; Pred. No. 0;
 Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY 1 ATGACCTGTGTAAGGCTCGCGGCACTTTGGGCAACAAGGCAAGACATATTATG 60
 DB 163 ATGACCTGTGTAAGGCTCGCGGCACTTTGGGCAACAAGGCAAGACATATTATG 222
 QY 61 AAATACGAAAGGAGCAGCGAGCTGGCTGCAGAGGCAAGGGGCTTAAGCTTTTGA 120
 DB 223 AAATACGAAAGGAGCAGCGAGCTGGCTGCAGAGGCAAGGGGCTTAAGCTTTTGA 282
 QY 121 AGCTCAACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTG 180
 DB 283 AGCTCAACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTG 342
 QY 181 ACTGCGGGGAGGAGCAAAATTCGGCGGAGATCAGCGGAAAGAGCAAGTGGTGGAT 240
 DB 343 ACTGCGGGGAGGAGCAAAATTCGGCGGAGATCAGCGGAAAGAGCAAGTGGTGGAT 402
 QY 241 ATGCTGGAGACTGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCTTAAAG 300
 DB 403 ATGCTGGAGACTGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCTTAAAG 462
 QY 301 GGAATGCGCATGAAATCCGCGGCGCCGATGTGTCAGTCTCTTGAACATTGAGAAATG 360
 DB 463 GGAATGCGCATGAAATCCGCGGCGCCGATGTGTCAGTCTCTTGAACATTGAGAAATG 522
 QY 361 AAGTTGAAAAAATCCCGGAAAGTACCAATCATGTAAGGAGAAAGGCAAGAGTCTGAG 420
 DB 523 AAGTTGAAAAAATCCCGGAAAGTACCAATCATGTAAGGAGAAAGGCAAGAGTCTGAG 582

QY 421 CACATCCAGCGCATTCGACCGGGAGCTTAAGCGGAGCATTAAGAGCATATATTTCTCAGG 480
 DB 583 CACATCCAGCGCATTCGACCGGGAGCTTAAGCGGAGCATTAAGAGCATATATTTCTCAGG 642
 QY 481 GATCGATACGGAACCAAGCAGCGGGAATTAATCTCAATCTCTCTGGATATGAGAGTAT 540
 DB 643 GATCGATACGGAACCAAGCAGCGGGAATTAATCTCAATCTCTCTGGATATGAGAGTAT 702
 QY 541 AACCCGGAGGTGGGCTAATCTGACAGGAGCTGAGCAATGCGGCTTTGTTCTCTCTAT 600
 DB 703 AACCCGGAGGTGGGCTAATCTGACAGGAGCTGAGCAATGCGGCTTTGTTCTCTCTAT 762
 QY 601 CTTCTGAGAGGATGATGATCTGAGGACTGTGTCAGCTGTCAGAGTGAAGAGCACTCC 660
 DB 763 CTTCTGAGAGGATGATGATCTGAGGACTGTGTCAGCTGTCAGAGTGAAGAGCACTCC 822
 QY 661 CTGACGAGATTTTCAAGACCCAAATGCGGAGACCTGTCAGGAGGCTCCAAAGCAAGAG 720
 DB 823 CTGACGAGATTTTCAAGACCCAAATGCGGAGACCTGTCAGGAGGCTCCAAAGCAAGAG 882
 QY 721 CATGTGTAGCCAGCTCAACCAAGACCATGAGGAGCATGAGCAAAAGATTCATATG 780
 DB 883 CATGTGTAGCCAGCTCAACCAAGACCATGAGGAGCATGAGCAAAAGATTCATATG 923
 QY 781 GGGCAGTGTCCCGTTAGGCTCTCATCCGATATTGATGACGGGATCTCTCTGGG 840
 DB 924 -----GATCTCTCTGGG 926
 QY 841 CTGACCTGTGCTGTGAGGAGCTGTATCTGTGTGAAGAGCAACAGGCTTATGATCCGAT 900
 DB 937 CTGACCTGTGCTGTGAGGAGCTGTATCTGTGTGAAGAGCAACAGGCTTATGATCCGAT 966
 QY 901 ACAAGATGCGCTTTAAGGTTCAGAGAGCGCTCAAGAGCGTCCAGGTGTGGCCG 960
 DB 997 ACAAGATGCGCTTTAAGGTTCAGAGAGCGCTCAAGAGCGTCCAGGTGTGGCCG 1056
 QY 961 TGGGCACTTTTTCGAACCGGTTCTGTGATCCTGTGGGCAAGGATGAGCACTGTGCTC 1020
 DB 1057 TGGGCACTTTTTCGAACCGGTTCTGTGATCCTGTGGGCAAGGATGAGCACTGTGCTC 1116
 QY 1021 AAGCATTTTGGGCTCTTATGAAGAACTTAACAAGAAAGAGGAGGACCTGCAACCCCA 1080
 DB 1117 AAGCATTTTGGGCTCTTATGAAGAACTTAACAAGAAAGAGGAGGACCTGCAACCCCA 1176
 QY 1081 GCCAAACCGAGGAAGGAGTGTGGGATCCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1140
 DB 1177 GCCAAACCGAGGAAGGAGTGTGGGATCCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1236
 QY 1141 ACCCTGTGCAAGGAGGAGCAGGAGCGCCCTCAGGCGCAACAGCCCGGTTCCCGGCGCC 1200
 DB 1237 ACCCTGTGCAAGGAGGAGCAGGAGCGCCCTCAGGCGCAACAGCCCGGTTCCCGGCGCC 1296
 QY 1201 ATTTGGTCAAGTTCCTGCGGCAAGGCACTGTTCTTCAACACCTGTCTGTGTGGGCT 1260
 DB 1297 ATTTGGTCAAGTTCCTGCGGCAAGGCACTGTTCTTCAACACCTGTCTGTGTGGGCT 1356
 QY 1261 GTCGGGGAAGACCTTAACCTGTGGGCACTCAGGAGTGTGCGGCGCCGCTGGGCTCAG 1320
 DB 1357 GTCGGGGAAGACCTTAACCTGTGGGCACTCAGGAGTGTGCGGCGCCGCTGGGCTCAG 1416
 QY 1321 GGAAGACCTCAGGAGTCTGAGAGATTCCTGAGTGAATCTCATGCCCCGCTTCCAAACG 1380
 DB 1417 GGAAGACCTCAGGAGTCTGAGAGATTCCTGAGTGAATCTCATGCCCCGCTTCCAAACG 1476
 QY 1381 GACCTGAGAGTGAAGGAGGCTTGTGTTCCGCTATTAATGATTTCAACAAGCTCTGGGCT 1440
 DB 1477 GACCTGAGAGTGAAGGAGGCTTGTGTTCCGCTATTAATGATTTCAACAAGCTCTGGGCT 1536
 QY 1441 CGTGCATATTCGAGAGAGACCAAGTGGCCCTGTCTGAGAGGCTGAACAACCTGTGGAG 1500
 DB 1537 CGTGCATATTCGAGAGAGACCAAGTGGCCCTGTCTGAGAGGCTGAACAACCTGTGGAG 1596

QY 1501 CGGGTAGATCGGCTTTCGCTGCAACCCAGACAGTGAATCCGACAGGACAGCCCTTCAGA 1560
 DB 1597 CGGGTAGATCGGCTTTCGCTGCAACCCAGACAGTGAATCCGACAGGACAGCCCTTCAGA 1556
 QY 1561 GCTAGGAGCAACAGCCGTGTCTCCCACTCAGAGGCTTTCCTCTCGGCTTCACCTTG 1620
 DB 1657 GCTAGGAGCAACAGCCGTGTCTCCCACTCAGAGGCTTTCCTCTCGGCTTCACCTTG 1716
 QY 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAGATCTGGGCAAGGCTCATGGCTGGAT 1680
 DB 1717 GAAAGTTCAGTTCCTCCAGGCTTCTAGAGATCTGGGCAAGGCTCATGGCTGGAT 1776
 QY 1681 AATTTCCCTAGGCTTAACAACCCAGCAAGCTTTCGCTCTCTTATTTTGTAA 1740
 DB 1777 AATTTCCCTAGGCTTAACAACCCAGCAAGCTTTCGCTCTCTTATTTTGTAA 1836
 QY 1741 CTTATGAAATGTATTTAAGAAAGTGCAGCTCGAGAGATTCAGAGATGAAACACACC 1800
 DB 1837 CTTATGAAATGTATTTAAGAAAGTGCAGCTCGAGAGATTCAGAGATGAAACACACC 1896
 QY 1801 AGACCCAGATCAACAAGCCAGATCCAGCCCTCCAGCAACCCCGCCAGAC 1860
 DB 1897 AGACCCAGATCAACAAGCCAGATCCAGCCCTCCAGCAACCCCGCCAGAC 1956
 QY 1861 CATGCTTGAATTCGACGACACCGTGAAGCTGCTTGTACTTTAACTCATGGAAG 1920
 DB 1957 CATGCTTGAATTCGACGACACCGTGAAGCTGCTTGTACTTTAACTCATGGAAG 2016
 QY 1921 ATAACTACCTTCAAGTTTGAATTAATGTTTCTGTGAAATG 1964
 DB 2017 ATAACTACCTTCAAGTTTGAATTAATGTTTCTGTGAAATG 2060

RESULT 14

AD44327

ID AD44327 standard; cDNA; 2072 BP.

XX AD44327;

DT 03-JUN-2004 (first entry)

DE Novel human arginine-rich protein cDNA #691.

KW ss; gene: human; arginine-rich protein; cancer; inflammation;

XX genetic disorder.

OS Homo sapiens.

PN US2004053250-A1.

PD 18-MAR-2004.

PF 21-NOV-2002; 2002US-00302172.

PR 05-MAR-2001; 2001US-00799451.

PR 05-MAR-2002; 2002US-00505095.

PR 20-AUG-2002; 2002US-00225251.

PA (TANG/) TANG Y T.

PA (XUEA/) XUE A.

PA (DRMA/) DRMANAC R T.

PI Tang YT, Xue A, Drmanac RT;

XX WPI; 2004-238579/22.

XX New isolated arginine-rich protein-like polynucleotides and polypeptides,

XX useful for diagnosing and/or treating conditions associated with aberrant

XX activity of the arginine-rich polypeptides, such as cancer and

XX inflammation.

CC The invention relates to an isolated polynucleotide. The methods and
 CC compositions of the present invention are useful for the diagnosis and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the arginine-rich protein-like polypeptides, such as
 CC cancer and inflammation. They can also be used in forensic, gene
 CC mapping, identification of mutations responsible for genetic disorders,
 CC and in assessing biodiversity. The present sequence represents a novel
 CC human arginine-rich protein cDNA.

SQ Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;

Query Match 92.3%; Score 1812.4; DB 12; Length 2072;

Best Local Similarity 96.3%; Pred. No. 0;

Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY 1 ATGACGTGTAGAGTTCGGGCGAGTGTGGGCAAGAGCGGAGGACATCATATG 60
 DB 163 ATGACGTGTAGAGTTCGGGCGAGTGTGGGCAAGAGCGGAGGACATCATATG 222
 QY 61 AATAGAAAAGGAGACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCA 120
 DB 223 AATAGAAAAGGAGACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCA 282
 QY 121 AGCTACAACAACAGCTGATCATTTGGGGAATGTACATGAGAGGAGTGCCTCTG 180
 DB 283 AGCTACAACAACAGCTGATCATTTGGGGAATGTACATGAGAGGAGTGCCTCTG 342
 QY 181 ACTGCGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240
 DB 343 ACTGCGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 402
 QY 241 ATGCTGGAGACTGGGAGAAATACAAACAGAGAAAGCTCATATGAGGCTACAG 300
 DB 403 ATGCTGGAGACTGGGAGAAATACAAACAGAGAAAGCTCATATGAGGCTACAG 462
 QY 301 GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTCTCTGAAACATGAGAAATG 360
 DB 463 GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTCTCTGAAACATGAGAAATG 522
 QY 361 AAGTTGAAAAAATCCCGGAAAGATACAGATCATGAAGAGAGGCAAGGTCATGTAG 420
 DB 523 AAGTTGAAAAAATCCCGGAAAGATACAGATCATGAAGAGAGGCAAGGTCATGTAG 582
 QY 421 CACATCCAGCGATGACCCGGGACGTAAACGGGACATTAAAGAAAGCATATTTCTCAG 480
 DB 583 CACATCCAGCGATGACCCGGGACGTAAACGGGACATTAAAGAAAGCATATTTCTCAG 642
 QY 481 GATGATACGGAACCAAGCAGGCGGAACTAATCTCAATCCTCTGCAATATGAGAGTAT 540
 DB 643 GATGATACGGAACCAAGCAGGCGGAACTAATCTCAATCCTCTGCAATATGAGAGTAT 702
 QY 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATGCGGCTTGTCTCTCTAT 600
 DB 703 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATGCGGCTTGTCTCTCTAT 762
 QY 601 CTTCTGAGAGAGATTCATTTGGGCACTGTGTCACTGTGCTGCTGAGAGGCACTCC 660
 DB 763 CTTCTGAGAGAGATTCATTTGGGCACTGTGTCACTGTGCTGCTGAGAGGCACTCC 822
 QY 661 CTGAGGAGATTCACAGCCCAATGGCGGAGCCGTCAAGGGGCTCCAAGCAACAGGAG 720
 DB 823 CTGAGGAGATTCACAGCCCAATGGCGGAGCCGTCAAGGGGCTCCAAGCAACAGGAG 882
 QY 721 CATGTGTAGCAAGTCAACCAACCAAGCATGAGGAGCATCAGGACAAAGAAATCATGT 780
 DB 883 CATGTGTAGCAAGTCAACCAACCAAGCATGAGGAGCATCAGGACAAAGAAATCATGT 923
 QY 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTTATTTAGCGGATCTCTCGGG 840
 DB 924 -----GATCTCTCTCGGG 936
 QY 841 CTCACCTGCGCTGTGGAGCGTATCTGTAGAAAGGCAAGGCGTTGATGCGGATA 900

```

Db      937 CTCAACCTGGGCGCTGTGGAGCTGTATCTGTAGTAGAGGCGAAGAGGCGTGTATGCCGATA 996
Qy      901 ACAAGATTCGCTTAAAGTTCAAGCAAGGCGCTCAAGAGAGCTCAGGTGNGCCG 960
Db      997 ACAAGATTCGCTTAAAGTTCAAGCAAGGCGCTCAAGAGAGCTCAGGTGNGCCG 1056
Qy      961 TGGGCAAGTTTTCGAAACCGGTTGATATCTGGGCGAAGGATGAGACACTGTGCTC 1020
Db      1057 TGGGCAAGTTTTCGAAACCGGTTGATATCTGGGCGAAGGATGAGACACTGTGCTC 1116
Qy      1021 AAGCATTTAAGGCTCTTAAGAAACTAAAGAAAGAGGAGGAGCTGCCACCCCA 1080
Db      1117 AAGCATTTAAGGCTCTTAAGAAACTAAAGAAAGAGGAGGAGCTGCCACCCCA 1176
Qy      1081 GCCAAACCCGAGCAAGGCTGTGGCATTCAGGCTGTGGCCGCTTACAGTGGGGAG 1140
Db      1177 GCCAAACCCGAGCAAGGCTGTGGCATTCAGGCTGTGGCCGCTTACAGTGGGGAG 1236
Qy      1141 ACCCTCTGCAAGGGGAGCAGGCAAGGCGCTTCAGAGCCCAAGGCGCGGTTCCCGGCGC 1200
Db      1237 ACCCTCTGCAAGGGGAGCAGGCAAGGCGCTTCAGAGCCCAAGGCGCGGTTCCCGGCGC 1296
Qy      1201 ATTTGTCAGCTTCCCGCAAGGCGCACTTCGTTCTTCACACCTGTCTGTGGGACT 1260
Db      1297 ATTTGTCAGCTTCCCGCAAGGCGCACTTCGTTCTTCACACCTGTCTGTGGGACT 1356
Qy      1261 GTCCGGGAAGACCTTACCTGTGGGAGCTCAGAGGCTGTGGCCGCGCTGGCTCAG 1320
Db      1357 GTCCGGGAAGACCTTACCTGTGGGAGCTCAGAGGCTGTGGCCGCGCTGGCTCAG 1416
Qy      1321 GGAAGACCTCAGAGGCTTCTGAGATTCCTGAGTGAACCTCATGCCCGCTCCCAAG 1380
Db      1417 GGAAGACCTCAGAGGCTTCTGAGATTCCTGAGTGAACCTCATGCCCGCTCCCAAG 1476
Qy      1381 GACCTGAGACGTAGAGGCGCTTGGTTCGCCATTATGATTTCAAGAGAGCTGTGGGCTC 1440
Db      1477 GACCTGAGACGTAGAGGCGCTTGGTTCGCCATTATGATTTCAAGAGAGCTGTGGGCTC 1536
Qy      1441 CGGCGCATATCCGAGAGGAGACAGCTGGCGCCCTGTGGGAGGCTGAACCTCGCGGAG 1500
Db      1537 CGGCGCATATCCGAGAGGAGACAGCTGGCGCCCTGTGGGAGGCTGAACCTCGCGGAG 1586
Qy      1501 CGGCTGAGATCGGCTTTCGCTGCAACCGACAGTATTCGACAGGAGACCCCTTCAGA 1560
Db      1597 CGGCTGAGATCGGCTTTCGCTGCAACCGACAGTATTCGACAGGAGACCCCTTCAGA 1656
Qy      1561 GCTAGGAGACGAACGCCGTGTGCTCCACTCAGAGGCTTGTGCTGCGCTTCACATTG 1620
Db      1657 GCTAGGAGACGAACGCCGTGTGCTCCACTCAGAGGCTTGTGCTGCGCTTCACATTG 1716
Qy      1621 GAAAGTCTCAGTTCCTCCGAGGCTTCTAAGAGATCTGGGCGAGGCTCATGCTGGAT 1680
Db      1717 GAAAGTCTCAGTTCCTCCGAGGCTTCTAAGAGATCTGGGCGAGGCTCATGCTGGAT 1776
Qy      1681 AATTTCCCTAAGGCTTAAACAACCAAGCAAGCTTCGCTCTGTTTTATTTTGGTTAA 1740
Db      1777 AATTTCCCTAAGGCTTAAACAACCAAGCAAGCTTCGCTCTGTTTTATTTTGGTTAA 1836
Qy      1741 CTTATGAAATGTATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800
Db      1837 CTTATGAAATGTATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1896
Qy      1801 AGACCCGAGATCAAAAGGCAACGATCCGAGCGCTCCGAGCAACCCGAGCCCAAGAC 1860
Db      1897 AGACCCGAGATCAAAAGGCAACGATCCGAGCGCTCCGAGCAACCCGAGCCCAAGAC 1956
Qy      1861 CATGCTCTGAATCTGACGACACCGTAGGCTGCTTGTGATCTTTAACTCATGAGAG 1920
Db      1957 CATGCTCTGAATCTGACGACACCGTAGGCTGCTTGTGATCTTTAACTCATGAGAG 2016
Qy      1921 ATAACTAATCTTACGTTTGAATTAATGTTTCTGTTGAATG 1964
Db      2017 ATAACTAATCTTACGTTTGAATTAATGTTTCTGTTGAATG 2060

```

```

RESULT 15
ADC37388
ID ADC37388 standard, DNA; 2647 BP.
XX
XX ADC37388;
AC
XX
XX 18-DEC-2003 (first entry)
DT
XX
XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 221.
DE
XX
XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KM cancer; infectious disease; bone disease; AIDS;
KM neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KM immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KM Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
XX NC02003048202-A2.
PN
XX
XX 12-JUN-2003.
PD
XX
XX 03-DEC-2002; 2002MO-JP012644.
PF
XX
XX 03-DEC-2001; 2001JP-00366692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002US-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH ) ASAH KASEI KK.
XX
XX Matsuda A, Muramatsu S;
PI
XX
XX WPI; 2003-505282/47.
DR P-PSDB; ADC37389.
XX
XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancer,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
XX Claim 4; SEQ ID NO 221; 938bp; English.
XX
XX The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
XX
XX Sequence 2647 BP; 608 A; 766 C; 750 G; 523 T; 0 U; 0 Other;
SQ
Query Match 92.1%; Score 1809; DB 10; Length 2647;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 5; Indels 66; Gaps 1;
Qy 1 ATGAGCTGTGTAGAGTTCGGGCGAGTGTGGGCGACAGAGGAGAGGACATCATTTATG 60
Db 755 ATGAGCTGTGTAGAGTTCGGGCGAGTGTGGGCGACAGAGGAGAGGACATCATTTATG 814
Qy 61 AAATPAGAAAGGAGACCGAGGCTGGGCTGCGAGAGGAGCAAGGGGCTTAAGCTTTTGA 120
Db 815 AAATPAGAAAGGAGACCGAGGCTGGGCTGCGAGAGGAGCAAGGGGCTTAAGCTTTTGA 874
Qy 121 AGCTACAACAACAGTCGATCATTTGGGATTTGATGATGAGCGAGCTGCTCTCTG 180
Db 875 AGCTACAACAACAGTCGATCATTTGGGATTTGATGATGAGCGAGGCTGCTCTCTG 934
Qy 181 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCATGCCGAAAGAGCAAGTGGTGGAT 240
Db 935 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCATGCCGAAAGAGCAAGTGGTGGAT 994

```

```

OY 241 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATATGAGCGTACAG 300
DB 995 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATATGAGCGTACAG 1054
OY 301 GGAATGCCATGAAACATCCGGGGCCGATGTGTGATGCTCTCTGAACATTAAGGAAATG 360
DB 1055 GGAATGCCATGAAACATCCGGGGCCGATGTGTGATGCTCTCTGAACATTAAGGAAATG 1114
OY 361 AAGTTGAAAACCCCGGAAAGATACAGATCATGAAGAGAGAGGCAAGGTCATTTGAG 420
DB 1115 AAGTTGAAAACCCCGGAAAGATACAGATCATGAAGAGAGAGGCAAGGTCATTTGAG 1174
OY 421 CACATCCAGCGCATGACCGGGACGTAAAGCGGGACATTAAGGAAAGATATTTCTCAG 480
DB 1175 CACATCCAGCGCATGACCGGGACGTAAAGCGGGACATTAAGGAAAGATATTTCTCAG 1234
OY 481 GATCGATACGGAACCAAGCAGCGGAACTACTCAATCTCTCTGATATGAGAGATAT 540
DB 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCAATCTCTCTGATATGAGAGATAT 1294
OY 541 AACCCGGAGTGGGCTACTGACGGAACCTGACCATGCGCGCTTGTCTCTCTAT 600
DB 1295 AACCCGGAGTGGGCTACTGACGGAACCTGACCATGCGCGCTTGTCTCTCTAT 1354
OY 601 CTTCCGAGGAGATGCAATTCGCGGACGTGAGCGTCTGCGCAGTGAAGAGGACCTCC 660
DB 1355 CTTCCGAGGAGATGCAATTCGCGGACGTGAGCGTCTGCGCAGTGAAGAGGACCTCC 1414
OY 661 CTGCGAGGATTTCAAGCCCAATGCGGGACCGTCAAGGGCTCCAGAACCAAGAGAG 720
DB 1415 CTGCGAGGATTTCAAGCCCAATGCGGGACCGTCAAGGGCTCCAGAACCAAGAGAG 1474
OY 721 CATGTGTAGCCAGCTCAACACCAAGCATGCGGCATCAAGACAAAGAAATCTATGT 780
DB 1475 CATGTGTAGCCAGCTCAACACCAAGCATGCGGCATCA----- 1515
OY 781 GGGGAGTGTCCCGCTGAGCGCTCATCCGGATATGATGACGGGATCTCTCCGG 840
DB 1516 -----GATCTCTCCGG 1528
OY 841 CTCAACCTGCGCTGTGAGACGTATCTGTATGAGAGGCAAGAGCGTTGATGCGATA 900
DB 1529 CTCAACCTGCGCTGTGAGAGCGTATCTGTATGAGAGGCAAGAGCGTTGATGCGATA 1588
OY 901 ACAAGATCGCTTTAAGTTTCAAGAAAGCGCTCAAGAGACGTCCAGGTGTGCGCG 960
DB 1589 ACAAGATCGCTTTAAGTTTCAAGAAAGCGCTCAAGAGACGTCCAGGTGTGCGCG 1648
OY 961 TGGGACGTTTTTGAACCGGTTGTTGATACCTGGGCAAGGATGAGAGACATGTGCTC 1020
DB 1649 TGGGACGTTTTTGAACCGGTTGTTGATACCTGGGCAAGGATGAGAGACATGTGCTC 1708
OY 1021 AAGCATCTTAAGGCTCTATGAAAGAACTAAACAAGAAAGAGGGGACCTGCAACCCCA 1080
DB 1709 AAGCATCTTAAGGCTCTATGAAAGAACTAAACAAGAAAGAGGGGACCTGCAACCCCA 1768
OY 1081 GCCAAACCCGAGCAAGGCTGTGCGCATCAAGGCTGTGCGGCTTCAAGTGGCGGAAAG 1140
DB 1769 GCCAAACCCGAGCAAGGCTGTGCGCATCAAGGCTGTGCGGCTTCAAGTGGCGGAAAG 1828
OY 1141 ACCCTGTCAGAGGGGAGAGAGGCGCTCAAGGCGCAAGAGCGGCTTCCCGGCGCC 1200
DB 1829 ACCCTGTCAGAGGGGAGAGAGGCGCTTCAAGGCGCAAGAGCGGCTTCCCGGCGCC 1888
OY 1201 ATTGTGTCAGCTTCCCGCACAAGGCACTGTTCTTCAACACCTGTCTGTGAGGCT 1260
DB 1889 ATTGTGTCAGCTTCCCGCACAAGGCACTGTTCTTCAACACCTGTCTGTGAGGCT 1948
OY 1261 GTCCGGGAAACACTTACCTGTGTGGCACTCAAGGCTGTGCCAGCCGGCCCTGCTCAG 1320
DB 1949 GTCCGGGAAACACTTACCTGTGTGGCACTCAAGGCTGTGCCAGCCGGCCCTGCTCAG 2008

```

```

OY 1321 GGAGGACCTCAGGTTTCCTGAGATTCCTGAGTGAATCCAGTGAATCCAGTCCCGGCTCCCAAG 1380
DB 2009 GGAGGACCTCAGGTTTCCTGAGATTCCTGAGTGAATCCAGTGAATCCAGTCCCGGCTCCCAAG 2068
OY 1381 GACCTGAGACGTAGAGGGCTTGTGTCGCGCATTAATGATTTTCAAGAGAGTGTGAGTTC 1440
DB 2069 GACCTGAGACGTAGAGGGCTTGTGTCGCGCATTAATGATTTTCAAGAGAGTGTGAGTTC 2128
OY 1441 CGTGCATATCCCAAGAGAGACCAAGTGGCCCCCTCTGTCAGAGGCTGAACACCTTGGAG 1500
DB 2129 CGTGCATATCCCAAGAGAGACCAAGTGGCCCCCTCTGTCAGAGGCTGAACACCTTGGAG 2188
OY 1501 CGGTCGATGAGTGGCTTGTGTCGTCACCAAGCACTGATTCGACACAGGGACCCCTTCAGA 1560
DB 2189 CGGTCGATGAGTGGCTTGTGTCGTCACCAAGCACTGATTCGACACAGGGACCCCTTCAGA 2248
OY 1561 GCTAGGAGAGAACAGCCGCTGTGTCACCACTCAAGGCGCTTGGCTGCGGCTCCACTTG 1620
DB 2249 GCTAGGAGAGAACAGCCGCTGTGTCACCACTCAAGGCGCTTGGCTGCGGCTCCACTTG 2308
OY 1621 GAAATTTCTCAGTTCCCTGACAGGCTTGAAGACATCTGGGCTCAGGGCTCATGCTGAT 1680
DB 2309 GAAATTTCTCAGTTCCCTGACAGGCTTGAAGACATCTGGGCTCAGGGCTCATGCTGAT 2368
OY 1681 AATTTCCCTAGGCTTGAACCAAGCAAGGCTTGGCTTGAATTTTGGTTAA 1740
DB 2369 AATTTCCCTAGGCTTGAACCAAGCAAGGCTTGGCTTGAATTTTGGTTAA 2428
OY 1741 CTTATGAAATATGATTAAGAAAGAGTGAAGTCAAGTCAAGAGATTCAGAGTGAACACACC 1800
DB 2429 CTTATGAAATATGATTAAGAAAGAGTGAAGTCAAGTCAAGAGATTCAGAGTGAACACACC 2488
OY 1801 AGACCCAGATCAAAAGCCAAACCATGCCCCAGCTTCCAGACACCCCAAGCCCAAGAC 1860
DB 2489 AGACCCAGATCAAAAGCCAAACCATGCCCCAGCTTCCAGACACCCCAAGCCCAAGAC 2548
OY 1861 CATGTTCTGAATTCGACGACACCGTAGGCTGCTTGAATTTAACTTAACTGAAGAG 1920
DB 2549 CATGTTCTGAATTCGACGACACCGTAGGCTGCTTGAATTTAACTTAACTGAAGAG 2608
OY 1921 ATTAACCTTCAAGTTTGAATTAATGTTTCTGTG 1959
DB 2609 ATTAACCTTCAAGTTTGAATTAATGTTTCTGTG 2647

```

Search completed: April 5, 2006, 13:32:41
Job time : 821 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:54:55 ; Search time 362 Seconds
(without alignments)
9644.009 Million cell updates/sec

Title: US-10-071-838-1
Perfect score: 1964
Sequence: 1 atcgacgtgtatagagtcgc.....aaatgttcctgttgaatcg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/prodata/1/ina/1/COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5/COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/6/COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/6/COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/6/COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/6/COMB.seq:*
 - 7: /cgn2_6/prodata/1/ina/6/COMB.seq:*
 - 8: /cgn2_6/prodata/1/ina/6/COMB.seq:*
 - 9: /cgn2_6/prodata/1/ina/6/COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836.4	93.5	2084	3	US-09-799-451-692
2	1812.4	92.3	2072	3	US-09-799-451-691
3	999	50.9	8201	2	US-08-253-155A-9
4	121	6.2	121	3	US-09-513-999C-24808
5	92	4.7	1925	3	US-09-949-016-2139
6	74.2	3.8	187	3	US-09-513-999C-13267
7	74	3.8	6453	2	US-08-306-691B-14
8	74	3.8	6453	3	US-09-209-668-10
9	74	3.8	6453	3	US-09-356-952-8
10	72.6	3.7	1981	3	US-09-620-312D-715
11	70	3.6	1823	3	US-09-774-528-336
12	70	3.6	1823	3	US-10-120-988-336
13	68.4	3.5	7218	2	US-08-232-463-14
14	63.4	3.2	806	3	US-09-870-767-9899
15	53.6	2.7	69199	3	US-09-949-016-13881
16	52.4	2.4	1887	3	US-09-270-767-12482
17	46.8	2.3	601	3	US-09-949-016-51412
18	45.4	2.3	25431	3	US-09-949-016-13334
19	45.4	2.3	2086	3	US-10-104-047-56
20	42.6	2.2	3039	3	US-09-620-312D-676
21	42.6	2.2	3039	3	US-09-620-312D-675
22	42	2.1	289	3	US-09-007-005-17
23	42	2.1	289	3	US-09-244-796-17
24	42	2.1	289	3	US-09-244-796-17

25	41.4	2.1	505	3	US-09-621-976-15639	Sequence 15639, A
26	40.4	2.1	1984	2	US-09-949-016-3047	Sequence 3047, Ap
27	40.2	2.0	13987	2	US-08-804-227C-13	Sequence 13, Appl
28	40.2	2.0	44377	2	US-08-804-227C-7	Sequence 7, Appl
29	40.2	2.0	44377	2	US-08-804-198-1	Sequence 1, Appl
30	40.2	2.0	53799	3	US-10-042-665A-3	Sequence 3, Appl
31	39.8	2.0	723	3	US-09-270-767-1491	Sequence 1491, Ap
32	39.8	2.0	723	3	US-09-270-767-16773	Sequence 16773, A
33	39.8	2.0	33478	3	US-09-949-016-13615	Sequence 13615, A
34	39	2.0	505	3	US-09-621-976-15639	Sequence 15639, A
35	38.6	2.0	482	3	US-09-392-184-8	Sequence 8, Appl
36	38.6	2.0	15722	3	US-09-949-016-16709	Sequence 16709, A
37	38.6	2.0	17000	3	US-09-548-797B-7	Sequence 7, Appl
38	38.6	2.0	17425	3	US-09-511-625B-5	Sequence 5, Appl
39	38.2	1.9	3983	3	US-09-762-311-3	Sequence 3, Appl
40	38.2	1.9	3983	3	US-09-762-311-4	Sequence 4, Appl
41	38	1.9	1179	3	US-10-077-176-71	Sequence 71, Appl
42	37.8	1.9	97423	3	US-09-949-016-12742	Sequence 12742, A
43	37.8	1.9	97424	3	US-09-949-016-15576	Sequence 15576, A
44	37.6	1.9	1056	3	US-09-902-540-5688	Sequence 5688, Ap
45	37.6	1.9	1152	3	US-09-902-540-2101	Sequence 2101, Ap

ALIGNMENTS

RESULT 1
US-09-799-451-692
Sequence 692, Application US/09799451
Percent No. 6783969

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing Ai
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungting
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE OF INVENTION: Polypeptides
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_flt_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (163) .. (1755)
US-09-799-451-692

Query Match 93.5%; Score 1836.4; DB 3; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

QY 1 ATCGACGTGTATAGAGTCGCGGCGAGTTGTGTGGGCAAGACGAGAGCATCTTANG 60
DB 163 ATCGACGTGTATAGAGTCGCGGCGAGTTGTGTGGGCAAGACGAGAGCATCTTANG 222

QY 61 AAATGCAAAAGGACACCGAGCTGGCTCCAGAGACAAAGGGGCTTACGCTTTTGA 120
 DB 223 AAAATACAAAAGGACACCGAGCTGGCTCCAGAGACAAAGGGGCTTACGCTTTTGA 282
 QY 121 AGCTACAAACAACAGCTGATCTTTGGGGATGTGACATGAGAGGAGCTGCTCTG 180
 DB 283 AGCTACAAACAACAGCTGATCTTTGGGGATGTGACATGAGAGGAGCTGCTCTG 342
 QY 181 ACTGCGGGAGGAGGAGCAAAATTCGGCGGAGATCAGCCGAAAGACAAATGGGTGAT 240
 DB 343 ACTGCGGGAGGAGGAGCAAAATTCGGCGGAGATCAGCCGAAAGAGAGATGGGTGAT 402
 QY 241 ATGCTGGAGACTGGAGAAATTAACAAAAGCAGAGAAAGCTCATGATCGAGCTACAG 300
 DB 403 ATGCTGGAGACTGGAGAAATTAACAAAAGCAGAGAAAGCTCATGATCGAGCTACAG 462
 QY 301 GGAATGCCATGAACTATCCGGGGCCGATGTGTCAGTCTCTGAACTTTGAGAAATG 360
 DB 463 GGAATGCCATGAACTATCCGGGGCCGATGTGTCAGTCTCTGAACTTTGAGAAATG 522
 QY 361 AAGTTGAAAAACCCGGAGATACAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 420
 DB 523 AAGTTG----- 529
 QY 421 CACATCCAGCGCATGACCGGAGCTGACCGGACATTAAGAGACATATATTTTCAG 480
 DB 530 -ACATCCAGCGCATGACCGGAGCTGACCGGACATTAAGAGACATATATTTTCAG 588
 QY 481 GATGATACGGAAACCAAGCAGCGGAACTACTCAATCTCTCTGCAATATGAGAAAT 540
 DB 589 GATGATACGGAAACCAAGCAGCGGAACTACTCAATCTCTCTGCAATATGAGAAAT 648
 QY 541 AACCCGAGGCTGCTACTGAGGAGCTGAGCCACATCGCGCTTGTCTCTCTAT 600
 DB 649 AACCCGAGGCTGCTACTGAGGAGCTGAGCCACATCGCGCTTGTCTCTCTAT 708
 QY 601 CTCTCTGAGAGATGATCTGGGCACTGGTGAAGCTGCTGGCAAGTGAAGGCACTCC 660
 DB 709 CTCTCTGAGAGATGATCTGGGCACTGGTGAAGCTGCTGGCAAGTGAAGGCACTCC 768
 QY 661 CTGAGAGGATTTTCACAGCCCAATGGCGGAGCTGCGAGGGGCTTCAAGACCAAGAG 720
 DB 769 CTGAGAGGATTTTCACAGCCCAATGGCGGAGCTGCGAGGGGCTTCAAGACCAAGAG 828
 QY 721 CATGTGATGACAGTCAACCAAGCACTGGGGCATCGAGCAAGAAAGATCTATGT 780
 DB 829 CATGTGATGACAGTCAACCAAGCACTGGGGCATCGAGCAAGAAAGATCTATGT 888
 QY 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTTGATTTGAAGGATCTCTCGGG 840
 DB 889 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTTGATTTGAAGGATCTCTCGGG 948
 QY 841 CTCACTGCGGCTGTGGAGCTGATCTGTGAGAGGCGAAACAGCGTTGATCCGATA 900
 DB 949 CTCACTGCGGCTGTGGAGCTGATCTGTGAGAGGCGAAACAGCGTTGATCCGATA 1008
 QY 901 ACAAGATCGCTTTAAGTTACAGCAGAGGCGCTCAAGAAAGCTCAGGCTGAGGCGG 960
 DB 1009 ACAAGATCGCTTTAAGTTACAGCAGAGGCGCTCAAGAAAGCTCAGGCTGAGGCGG 1068
 QY 961 TGGGCACTTTTTCACACCGGTTGTTGATCTGAGGCGAGGATGAGACACTGTGCTC 1020
 DB 1069 TGGGCACTTTTTCACACCGGTTGTTGATCTGAGGCGAGGATGAGACACTGTGCTC 1128
 QY 1021 AAGCATTTAAGGCTCTTATGAGAAATTAACAAAGAAAGCAGGGGACCTTGCAACCCCA 1080
 DB 1129 AAGCATTTAAGGCTCTTATGAGAAATTAACAAAGAAAGCAGGGGACCTTGCAACCCCA 1188
 QY 1081 GCCAAACCGAGAGAGGTGCTGCGCATCCAGGCGCTGCGGCTTACAGTGGCGGAGAG 1140
 DB 1189 GCCAAACCGAGAGAGGTGCTGCGCATCCAGGCGCTGCGGCTTACAGTGGCGGAGAG 1248
 QY 1141 ACCCTTGCAAGGGGAGCAGGCAAGGCCCTTCAAGGCCACCAAGCCCGGTTCCCGGCGC 1200

DB 1249 ACCCTTGCAAGGGGAGCAGGCAAGGCCCTTCAAGGCCACCAAGCCCGGTTCCCGGCGC 1308
 QY 1201 ATTTGTGACCTTCCCGGACAGGAGCACTGTTCTTTCACACCTGTCTGTGGGGCT 1260
 DB 1309 ATTTGTGACCTTCCCGGACAGGAGCACTGTTCTTTCACACCTGTCTGTGGGGCT 1368
 QY 1261 GTCGGGAGAGACCTACCTGTGGGCACTCAGGGGTGCGCAGCCCGGCTTGGCTCAG 1320
 DB 1369 GTCGGGAGAGACCTACCTGTGGGCACTCAGGGGTGCGCAGCCCGGCTTGGCTCAG 1428
 QY 1321 GAGAGACTTACGGGTTCTTGAAGTTCTTCACTGAGAACTTCATGCCCCCTTCCCAAG 1380
 DB 1429 GAGAGACTTACGGGTTCTTGAAGTTCTTCACTGAGAACTTCATGCCCCCTTCCCAAG 1488
 QY 1381 GACCTGACGTAGAGGAGCCCTTGGTTCCCGCATTTATGATTTTCAAGCAGAGCTGAGGTC 1440
 DB 1489 GACCTGACGTAGAGGAGCCCTTGGTTCCCGCATTTATGATTTTCAAGCAGAGCTGAGGTC 1548
 QY 1441 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGCTGAGGAGCTGAAACCCCTGCGAG 1500
 DB 1549 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGCTGAGGAGCTGAAACCCCTGCGAG 1608
 QY 1501 CGGATGATCGGCTTTCGCTGACCCAGCATGATTTCCGACAGAGGACCCCTTCAGA 1560
 DB 1609 CGGATGATCGGCTTTCGCTGACCCAGCATGATTTCCGACAGAGGACCCCTTCAGA 1668
 QY 1561 GCTAGGAGAGACAGCGGTGCTCCCACTCAGAGGCTTGGCCCTGCGGAGCTCACTTG 1620
 DB 1669 GCTAGGAGAGACAGCGGTGCTCCCACTCAGAGGCTTGGCCCTGCGGAGCTCACTTG 1728
 QY 1621 GAAAGTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGAGGCAAGGCTCATGAGTGTGAT 1680
 DB 1729 GAAAGTCTCAGTTCCCTCCAGGCTTCTAGAGCATCTGAGGCAAGGCTCATGAGTGTGAT 1788
 QY 1681 AATTTCCCTAGGCTTTCACACCCAGCAAGCTTGGCTCTGTTTATTTTGGTTAA 1740
 DB 1789 AATTTCCCTAGGCTTTCACACCCAGCAAGCTTGGCTCTGTTTATTTTGGTTAA 1848
 QY 1741 CTTATGAAATGTATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800
 DB 1849 CTTATGAAATGTATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1908
 QY 1801 AGACCCAGATCAACAAAGCCACATGCCAGCCCTCCAGAGACCCCGCCAGCCAGAC 1860
 DB 1909 AGACCCAGATCAACAAAGCCACATGCCAGCCCTCCAGAGACCCCGCCAGCCAGAC 1968
 QY 1861 CATGCTTGAATTTGACGACACCGTGAAGCTGCTTGTATCTTTAACTCATGGAAG 1920
 DB 1969 CATGCTTGAATTTGACGACACCGTGAAGCTGCTTGTATCTTTAACTCATGGAAG 2028
 QY 1921 ATTAATCACTTCACTTTGAATTAATGTTTCTGTGAAATG 1964
 DB 2029 ATTAATCACTTCACTTTGAATTAATGTTTCTGTGAAATG 2072

RESULT 2
 US-09-799-451-691
 ; Sequence 691, Application US/09799451
 ; Patent No. 6783969
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Ma, Yungting
 ; APPLICANT: Yamazaki, Victoria
 ; APPLICANT: Chen, Rui-hong

APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 691
LENGTH: 2072
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163)..(1743)
US-09-799-451-691

Query Match 92.3%; Score 1812.4; DB 3; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
QY 1 ATGAGCTGTGTAGAGTCCCGGCACTTGTGGCACAAGAGGAGAGACATCATTTAG 60
DB 163 ATGAGCTGTGTAGAGTCCCGGCACTTGTGGCACAAGAGGAGAGACATCATTTAG 222
QY 61 AATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAGGGGCTTAAGCTTTTCA 120
DB 223 AATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAGGGGCTTAAGCTTTTCA 282
QY 121 AGCTACAAACAACGCTGCATTTGGGATTTGATCATAGACGAGAGCTGCTCTCTG 180
DB 283 AGCTACAAACAACGCTGCATTTGGGATTTGATCATAGACGAGAGCTGCTCTCTG 342
QY 181 ACTGGCGGGAGGCGGAAGCAATTCCGCGGAGATAGCCGAAAGCAAGTGGTGGAT 240
DB 343 ACTGGCGGGAGGCGGAAGCAATTCCGCGGAGATAGCCGAAAGCAAGTGGTGGAT 402
QY 241 ATGCTGGGAGACGGGGAATTAACAAGACAGAAAGCTCATGATGAGAGGTACAG 300
DB 403 ATGCTGGGAGACGGGGAATTAACAAGACAGAAAGCTCATGATGAGAGGTACAG 462
QY 301 GGAATGCCCATGAATCCGCGGCGCGATGTGTCAGTCTCTGAACTTGAAGAAATG 360
DB 463 GGAATGCCCATGAATCCGCGGCGCGATGTGTCAGTCTCTGAACTTGAAGAAATG 522
QY 361 AAGTTGAAAAACCCGGAAGATACAGATCATGAAGAGAGGGCAAGAGTCTATGAG 420
DB 523 AAGTTGAAAAACCCGGAAGATACAGATCATGAAGAGAGGGCAAGAGTCTATGAG 582
QY 421 CACATCCAGCGCATGACCGGGAAGCTTAAGCGGACATTAAGAAAGATATTTTCAAG 480
DB 583 CACATCCAGCGCATGACCGGGAAGCTTAAGCGGACATTAAGAAAGATATTTTCAAG 642
QY 481 GATCATACGGAACCAAGAGCGGGAATCACTCAATCCTCTGCAATATGAGAGAT 540
DB 643 GATCATACGGAACCAAGAGCGGGAATCACTCAATCCTCTGCAATATGAGAGAT 702
QY 541 AACCCGAGAGTGGGCTACTGCAAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT 600
DB 703 AACCCGAGAGTGGGCTACTGCAAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT 762
QY 601 CTTCTGAGGAGATGATCTGGGCACTGTGCAAGTCTGCGCAATGAGAGGACATCC 660
DB 763 CTTCTGAGGAGATGATCTGGGCACTGTGCAAGTCTGCGCAATGAGAGGACATCC 822
QY 661 CTGACGAGATTTCAACGCGCAATGGCGGAGCCGTCTCAAGGAGCTCCAAGCAAGAG 720
DB 823 CTGACGAGATTTCAACGCGCAATGGCGGAGCCGTCTCAAGGAGCTCCAAGCAAGAG 882

QY 721 CATGTGTAGCAACGTCACAACCAAGACATTTGGGCAATCAGAGACAAGAAATCTATCT 780
DB 883 CATGTGTAGCAACGTCACAACCAAGACATTTGGGCAATCAGAGACAAGAAATCTATCT 923
QY 781 GGGCAGTGTCCCGGTAGGCTGCTCATCCGGAATTTGATGACGGGATCTCTCGGG 840
DB 924 -----GATCTCTCTCGGG 936
QY 841 CTCACCTTCGCGCTGTGGACGTGTATCTGTAGAAAGGGAACAGGCTTGTATCCGATTA 900
DB 937 CTCACCTTCGCGCTGTGGACGTGTATCTGTAGAAAGGGAACAGGCTTGTATCCGATTA 996
QY 901 ACAAGATTCCTTTTAAGTTTCAAGAAAGCCCTCACGAACGTCACAGTGTGGCCCG 960
DB 997 ACAAGATTCCTTTTAAGTTTCAAGAAAGCCCTCACGAACGTCACAGTGTGGCCCG 1056
QY 961 TGGGACGTTTTTGGCAACCGGTTGTTGATACCTGAGGCGAAGGATGAGGACATGTGCTC 1020
DB 1057 TGGGACGTTTTTGGCAACCGGTTGTTGATACCTGAGGCGAAGGATGAGGACATGTGCTC 1116
QY 1021 AAGCATCTTTAGGCGCTTTATGAAAGAACTTAACAAGAAAGCAAGGAGGACCTGACCCCA 1080
DB 1117 AAGCATCTTTAGGCGCTTTATGAAAGAACTTAACAAGAAAGCAAGGAGGACCTGACCCCA 1176
QY 1081 GCCAAACCCGACCAAGGCTGTGCGCATTCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1140
DB 1177 GCCAAACCCGACCAAGGCTGTGCGCATTCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1236
QY 1141 ACCCTCTGCAAGGAGGACAGGAGGCGCCCTCCAGGCGCAAGCCGAGTTCGCGAGGCC 1200
DB 1237 ACCCTCTGCAAGGAGGACAGGAGGCGCCCTCCAGGCGCAAGCCGAGTTCGCGAGGCC 1296
QY 1201 ATTTGTCAGCTTCCCGCAAGGCAAGCTGCTTTCACAACCTGTCTGTGTGGAGCT 1260
DB 1297 ATTTGTCAGCTTCCCGCAAGGCAAGCTGCTTTCACAACCTGTCTGTGTGGAGCT 1356
QY 1261 GTCCGGGAAGACACTACCTGTGGGCACTAGGAGTGTGCCAGGCCCGGCTGCTCAG 1320
DB 1357 GTCCGGGAAGACACTACCTGTGGGCACTAGGAGTGTGCCAGGCCCGGCTGCTCAG 1416
QY 1321 GGAAGACCTCAGGCTTCCGAGGATTCGAGAGTGAAGTCCATGACCCCGCTCCCAAG 1380
DB 1417 GGAAGACCTCAGGCTTCCGAGGATTCGAGAGTGAAGTCCATGACCCCGCTCCCAAG 1476
QY 1381 GACCTGACCTGAGAGGCGCTTGTGTCGCAATTAATGATTTTCAACAAGAGCTGTGGCTC 1440
DB 1477 GACCTGACCTGAGAGGCGCTTGTGTCGCAATTAATGATTTTCAACAAGAGCTGTGGCTC 1536
QY 1441 CGTGCATATCCCAAGAGGACCAAGTGGCGCTGCTGCGAGGCTGAACCTTGGGAG 1500
DB 1537 CGTGCATATCCCAAGAGGACCAAGTGGCGCTGCTGCGAGGCTGAACCTTGGGAG 1596
QY 1501 CGGAGGAGATGGGCTTGGCTGCAACCGACCTGATTCGACAGAGGACCCGCTTCAAG 1560
DB 1597 CGGAGGAGATGGGCTTGGCTGCAACCGACCTGATTCGACAGAGGACCCGCTTCAAG 1656
QY 1561 GCTAAGGACGAACAGCCGCTGTGCTCCACCTTCAAGGCGCTTGGCGCTTCAAGTGG 1620
DB 1657 GCTAAGGACGAACAGCCGCTGTGCTCCACCTTCAAGGCGCTTGGCGCTTCAAGTGG 1716
QY 1621 GAAAGTTCTCAGTTCCTTCAGAGCTTGAAGACATCTGGGCGAGGCTCATGGCTGAT 1680
DB 1717 GAAAGTTCTCAGTTCCTTCAGAGCTTGAAGACATCTGGGCGAGGCTCATGGCTGAT 1776
QY 1681 AATTTCCCTAAGGCTTAACAACCAAGAGCTTGGCTGCTTATTTTGGTTAA 1740
DB 1777 AATTTCCCTAAGGCTTAACAACCAAGAGCTTGGCTGCTTATTTTGGTTAA 1836
QY 1741 CTTATGAAATGTATTTAAGAAAGAGTGCAGCTCGAAGAGATTCAGAGATGGAACAACC 1800
DB 1837 CTTATGAAATGTATTTAAGAAAGAGTGCAGCTCGAAGAGATTCAGAGATGGAACAACC 1896

QY 1801 AGACCCAGATCACAAAGCCCAACCATGCGCCCTCCAGCAACCCCAAGCCCAAGAC 1860
DB 1897 AGACCCAGATCACAAAGCCCAACCATGCGCCCTCCAGCAACCCCAAGCCCAAGAC 1956
QY 1861 CATGTTCTGAATTTGACGACACCGTGAAGCTTCTTGTACTTTAACTCATGGAAG 1920
DB 1957 CATGTTCTGAATTTGACGACACCGTGAAGCTTCTTGTACTTTAACTCATGGAAG 2016
QY 1921 ATAATCACTTCACGTTTGAATTAATGTTTCTGTTGAATG 1964
DB 2017 ATAACCACTTCATGTTTGAATTAATGTTTCTGTTGAATG 2060

RESULT 3
US-08-253-155A-9
Sequence 9, Application US/08253155A
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jeno
APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIT-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 8201 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-253-155A-9

Query Match 50.9%; Score 999; DB 2; Length 8201;
Best Local Similarity 77.8%; Pred. No. 1,2e+263;
Matches 1332; Conservative 0; Mismatches 165; Indels 214; Gaps 2;

QY 241 ATCTGGAGATCTGGAGAAATACAAAGCAGCAGAAAGCTATAGATCGAGCTACAG 300
DB 1934 ATCTGGAGAAATGGGAGACATATTAAGCACAGTACGAATCTCATAGTCAGAGTACAG 1993
QY 301 GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTCTCTGAAACATTGAGAAATG 360
DB 1994 GGAATTCCTCAATGAACATCCGGGGCCGATGTGGTCACTCTCTGAAACATTGAGAAATC 2053
QY 361 AAGTTGAAAAACCCCGGAGATACAGATCATGAAGAGAGGAGGAGGATCATGAG 420
DB 2054 AAGTTGAAAAACCCCGGAGATACAGATCATGAAGAGAGGAGGAGGATCATGAG 2113
QY 421 CACATCCAGGCATTCGACCCGGAGCTTAAGCGGAGCATTAAGAGCATATATTTCTCAG 480
DB 2114 CACATCCAGCATTCGACCTGAGCGTGAAGAGACACTCCGGAACCATGCTTTTAGG 2173
QY 481 GATCATACGGAACCAAGCAGCGGAACTACTCAATCTCTGCGATATGAGAGTAT 540
DB 2174 GATCATATGGAACCAAGCAGAGGAACTAATTTACATCTCTCTGCGATATGAGAGTAT 2233
QY 541 AACCCGAGGTGGGCTAATGCAAGGACCTGAGCAGATGCGCGCTTGTCTCTCTAT 600
DB 2234 AACCCGAGGTGGGCTAATGCAAGGACCTGAGCAGATGCGCGCTTGTCTCTCTAT 2293
QY 601 CTTCCTGAGAGGATCATTTCTGGGCACTGTGCACTGCGCAGTGAAGCACTCC 660
DB 2294 CTGCTGAGAGGAGGACGATTTCTGGGCACTGTGCACTGCGCAGTGAAGCACTCC 2353
QY 661 CTGCAAGGATTTTACAGACCCCAATATGCGGGAACCTGTCAGGGGCTCCAAAGCAACGAG 720
DB 2354 CTGCAAGGATTTTACAGACCCCAATATGCGGGAACCTGTCAGGGGCTCCAAAGCAACGAG 2413
QY 721 CATGTGTAGCGACGTCACAAACCAAGACCATAGGGGATCAGGACAAAGATCTATGT 780
DB 2414 CATGTGTAGCGACGTCACAAACCAAGACCATAGGGGATCAGGACAAAGATCTATGT 2473
QY 781 GGGCAGTGTCTCCCGTTAGGCTGCTCATCCGATTTTGAATGACGCGATCTCTCGGG 840
DB 2474 GGGCAGTGTCTCCCGTTAGGCTGCTCATCCGATTTTGAATGACGCGATCTCTCGGG 2533
QY 841 CTCACTCTGCGCTGTGGGACGTGTATCTGTGTGAAGGCGAACAAGCGCTTATGCCGAT 900
DB 2534 CTCACTCTGCGCTGTGGGACGTGTATCTGTGTGAAGGCGAACAAGCGCTTATGCCGAT 2593
QY 901 ACAAGATGCTTTAAGTTGAGCAAGAGCGCTCAGGAAGCGTCCAGGATGAGGCGG 960
DB 2594 ACAAGATGCTTTAAGTTGAGCAAGAGCGCTCAGGAAGCGTCCAGGATGAGGCGG 2653
QY 961 TGGGCACTTTTTCACACCGGTTCTGTGATACCTGGGCGCAGGATGAGCACTGTGCTC 1020
DB 2654 TGGGCACTTTTTCACACCGGTTCTGTGATACCTGGGCGCAGGATGAGCACTGTGCTC 2713
QY 1021 AAGCATCTTAGGCGCTCTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCCACCC-- 1078
DB 2714 AAGCATCTTAGGCGCTCTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCCACCCCA 2773
QY 1079 ----- 1078
DB 2774 GGGCCAAACGCTTGGAGCAAGAGTGTGAGCAAGAACCCCAAGCAATCTGAACCTTG 2833
QY 1079 ----- 1078
DB 2834 GGGGCACTCCAGAGACCAACCAATGCCCCAAGGGCTTCCCATGCGACGACACAC 2893
QY 1079 ----- 1078
DB 2894 ACCCTCTCTTGGATTCAGACAGCTAAGGCGTGTCTAGTGTGACACAGAGGGCC 2953
QY 1079 -----CAGCAAAACCGAGCAAGGGTGTGCGGATC 1109
DB 2954 ACAAGAGACCCCAAGGACTTCAAGATGCAAGCAACGGAACAAAGGGTCTTGGAC 3013

Thu Apr 6 16:21:14 2006

us-10-071-838-1.rn1

```

QY 1110 CAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGAGACGAGCGCC 1169
DB 3014 CAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3073
QY 1170 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 1229
DB 3074 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3133
QY 1230 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 1289
DB 3134 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3193
QY 1290 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 1349
DB 3194 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3253
QY 1350 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 1409
DB 3254 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3313
QY 1410 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 1469
DB 3314 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3373
QY 1470 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 1500
DB 3374 TCAGGCTGTGCTCCGCTTCACTGTGCGGGAAGACCTCTGTGAGGGGATATGAGGCCCC 3404

```

RESULT 4

```

US-09-513-999C-24808
; Sequence 24808, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122, 487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 24808
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-24808

```

```

Query Match 6.2%; Score 121; DB 3; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.6e-23;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1737 TAACTTATGAATATATTAAGAAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACA 1796
DB 1 TAACTTATGAATATATTAAGAAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACA 60
QY 1797 CACGAGACCCGATCACAAGGCAATGCGGAGCCCTCCAGACCCCGAGCCCA 1856
DB 61 CACGAGACCCGATCACAAGGCAATGCGGAGCCCTCCAGACCCCGAGCCCA 120
QY 1857 C 1857
DB 121 C 121

```

```

RESULT 5
US-09-949-016-2139
; Sequence 2139, Application US/09949016
; Patent No. 6812339

```

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2139
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2139

```

```

Query Match 4.7%; Score 92; DB 3; Length 1925;
Best Local Similarity 60.3%; Pred. No. 8.1e-15;
Matches 152; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

```

QY 422 ACATCCAGGCAATCCAGCGGAGCTGAGCGGAGCATTAAGAGCATATATTTTCAGGG 481
DB 118 ACATCCAGGCAATCCAGCGGAGCTGAGCGGAGCATTAAGAGCATATATTTTCAGGG 177
QY 482 ATGATACGAGCAACCAAGCGGGAACCTACTCCATCTCTGCGATATGAGGATATA 541
DB 178 ACAGATATGATGTTAAGCAACATCTTATTCATGATGCTTGTCTTATTTATTA 237
QY 542 ACCCGAGGTGGCTTACTGAGGAGCTGAGCGGAGCTGAGCGGAGCATTAAGAGCATAT 601
DB 238 ACACGGAAGTGGGATATGCTGAGGAGATGAGCGGAGCATTAAGAGCATATTA 297
QY 602 TTCTGAGAGAGATGATCTTGGGAGCTGAGCGGAGCTGAGCGGAGCATTAAGAGCAT 661
DB 298 TGAACGAGAGATGATCTTGGGAGCTGAGCGGAGCTGAGCGGAGCATTAAGAGCAT 357
QY 662 TGCAGGATTTT 673
DB 358 TGCAGGATTTT 369

```

RESULT 6

```

US-09-513-999C-13267
; Sequence 13267, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122, 487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13267
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; LOCATION: 5
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature

```

LOCATION: 132
OTHER INFORMATION: y=c or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: 136
OTHER INFORMATION: y=c or t
US-09-513-999C-13267

Query Match 3.8%; Score 74.2; DB 3; Length 187;
Best Local Similarity 97.3%; Pred. No. 2.2e-10;
Matches 73; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCTGTGTAGAGGTGCGCGGCACTTGTGGCACAAGAGCGAGACATCATTTATG 60
DB 113 ATGACCTGTGTAGAGGTGCGGGAAGTTGTGGCACAAGAGCGAGACATCATTTATG 172

QY 61 AATACGAAAAAGGA 75
DB 173 AATACGAAAAAGGA 187

RESULT 7

US-08-306-691B-14/C
Sequence 14, Application US/08306691B
Patent No. 5734039
GENERAL INFORMATION:
APPLICANT: Calabretta, Bruno
APPLICANT: Skorski, Tomasz
TITLE OF INVENTION: ANTISENSE
TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavoroga & Monaco, P.C.
STREET: Two Penn Center, Suite 1800
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION NUMBER:
APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: NO. 5734039E
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6453 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-306-691B-14

Query Match 3.8%; Score 74; DB 2; Length 6453;
Best Local Similarity 58.9%; Pred. No. 1.2e-09;
Matches 166; Conservative 0; Mismatches 110; Indels 6; Gaps 2;

QY 1103 CGGCATCCAGGCTGTGCTCCGCGCTTCACGTGCGGGAAGACCTCTTGCAGGGGAGCAGGC 1162
DB 6200 CGGCATCCAGGCTGTGCTCCGCGGTAACCCACGACATCAAGACCTCTTCAAGGGGAGGAGC 6141

QY 1163 AGCCCCCTCAGAGCCCAAGCCCGGTTCCCGGCGCAATTGTCAGCTTCCCGGCAC 1222
DB 6140 TGGCTCTCCCTCCATGAGCAGAGCTCCAGAAATCCAAAGGGGTAGCT---CCCAAT 6084
QY 1223 GGGACCTGCTTTTTCACACCTGTCTGTGTGGGCTGTCCGGGAAGACACTTACCCTG 1282
DB 6083 GGGGGCTGTGCTTCCGACACACTGGCTTAGCTGGGCTGTCCCGAGG---CCAAAGGCT 6027
QY 1283 TGGGCACTGAGGGTGTGCCAGCCCGGCTGACTCAGGAGGACCTCAGGTTCTTGA 1342
DB 6026 AGCGGATGAGGGGCTTCCGTGAGAGCCAGGCAAGTGGGCTCTCTTGAAGTCCCTAG 5967
QY 1343 GATTCTGCAGTGAATCCATGCCCCGCTCCCAACGAGC 1384
DB 5966 GACTGGGGGTGTGGCTGGCTGTCCCAATGAGACAGACC 5925

RESULT 8

US-09-209-668-10/C
Sequence 10, Application US/09209668A
Patent No. 6114517
GENERAL INFORMATION:
APPLICANT: Monia, Brett P.
APPLICANT: Xu, Xiaoxing S.
TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
TITLE OF INVENTION: ALPHA-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
FILE REFERENCE: ISPH-0336
CURRENT APPLICATION NUMBER: US/09/209,668A
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 6453
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1664)..(1774)
FEATURE:
NAME/KEY: CDS
LOCATION: (2042)..(2220)
FEATURE:
NAME/KEY: CDS
LOCATION: (2374)..(2533)
FEATURE:
NAME/KEY: CDS
LOCATION: (3231)..(3350)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: V00574/Genbank
DATABASE ENTRY DATE: 1991-01-03
US-09-209-668-10

Query Match 3.8%; Score 74; DB 3; Length 6453;
Best Local Similarity 58.9%; Pred. No. 1.2e-09;
Matches 166; Conservative 0; Mismatches 110; Indels 6; Gaps 2;

QY 1103 CGGCATCCAGGCTGTGCTCCGCGCTTCACGTGCGGGAAGACCTCTTGCAGGGGAGCAGGC 1162
DB 6200 CGGCATCCAGGCTGTGCTCCGCGGTAACCCACGACATCAAGACCTCTTCAAGGGGAGGAGC 6141
QY 1163 AGCCCCCTCAGAGCCCAAGCCCGGTTCCCGGCGCAATTGTCAGCTTCCCGGCAC 1222
DB 6140 TGGCTCTCCCTCCATGAGCAGAGCTCCAGAAATCCAAAGGGGTAGCT---CCCAAT 6084
QY 1223 GGGACCTGCTTTTTCACACCTGTCTGTGTGGGCTGTCCGGGAAGACACTTACCCTG 1282
DB 6083 GGGGGCTGTGCTTCCGACACACTGGCTTAGCTGGGCTGTCCCGAGG---CCAAAGGCT 6027
QY 1283 TGGGCACTGAGGGTGTGCCAGCCCGGCTGACTCAGGAGGACCTCAGGTTCTTGA 1342
DB 6026 AGCGGATGAGGGGCTTCCGTGAGAGCCAGGCAAGTGGGCTCTCTTGAAGTCCCTAG 5967

Thu Apr 6 16:21:14 2006

us-10-071-838-1.rn1

Page 7

QY 1343 GATTCTGACGTGGAAGTCCATGCCCCCGCTCCCAACGAGACC 1384
Db 5966 GACTCGGGGTGTGGCTGGCTGTCTCCCAATCAGAGACACACC 5925

RESULT 9
US-09-356-952-8/c

/ Sequence 8, Application US/09356952
/ Patent No. 6117663
/ GENERAL INFORMATION:
/ APPLICANT: Boriack-Sjodin, Ann
/ APPLICANT: Margalit, S. M.
/ APPLICANT: Bor-Sogil, Dafna
/ APPLICANT: Cole, Philip
/ APPLICANT: Kurlyan, John
/ TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 600-1-228N
/ CURRENT APPLICATION NUMBER: US/09/356,952
/ CURRENT FILING DATE: 1999-07-19
/ EARLIER APPLICATION NUMBER: 60/093,631
/ EARLIER FILING DATE: 1998-07-21
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 8
/ LENGTH: 6453
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-356-952-8

Query Match 3.8%; Score 74; DB 3; Length 6453;
Best Local Similarity 58.9%; Pred. No. 1,2e-09;
Matches 166; Conservative 0; Mismatches 110; Indels 6; Gaps 2;

QY 1103 CGGCATTCAGAGCTGTGCGGGCTTCACTGTGGCGGAAAGACCTCTGCAAGGGGAGACGCC 1162
Db 6200 CGGCACCCAGAGCTGTGCGGGTAAACCCAGGCATCAAGACCTCTCCCAAGGGGAGAGC 6141
QY 1163 AGGCCCCCTCAGGCGCCAGACCCCGGTTCCCGGGCCCATTTGTGAGCTTCCCGGCAC 1222
Db 6140 TGGCTCTCCCTCCATCAGACAGCTCCAGATCCCAAGGGGTACGCT---CCACAT 6084
QY 1223 GGGCAGCTGCTTCTTCCACACACCTGTCTGTGGGGCTGTCGGGAAAGACCTACCTG 1282
Db 6083 GGGGGGCTGTCTCTGCAACACACTGGCTGTGCTGGGCTGTCTCCGAGG---CCAAGGCT 6027
QY 1283 TGGGCACTCAGGGGTGTGCCAGCCCGGCTTGTGCTCAAGGAGACCTCAAGGGTTCCTGA 1342
Db 6026 AGCGGATGAGGGGGCTTCCGTGAGGCCAGGGCAGTGGGCTGCTCTGTGGAAGTCCCTAG 5967
QY 1343 GATTCTGACGTGGAAGTCCATGCCCCCGCTCCCAACGAGACC 1384
Db 5966 GACTCGGGGTGTGGCTGGCTGTCTCCCAATCAGAGACACACC 5925

RESULT 10

US-09-620-312D-715

/ Sequence 715, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yongheng
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yuning

/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghaast
/ APPLICANT: Dramanc, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP28
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_FL_gene Version 1.0
/ SEQ ID NO 715
/ LENGTH: 1981
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (59)..(1585)
US-09-620-312D-715

Query Match 3.7%; Score 72.6; DB 3; Length 1981;
Best Local Similarity 50.5%; Pred. No. 1.7e-09;
Matches 233; Conservative 0; Mismatches 219; Indels 9; Gaps 2;

QY 184 GCGCGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGTATG 243
Db 266 GCGCTGAGAGAGTACCCCTCGAGGTGTGAGCAGAGGAGTCAAGTGGCTGACATG 325
QY 244 CTGGAGACTGGGAGAAAT-----ACAAAGCAGCAAGAACTCATATGACGCTAC 297
Db 326 CTCACACACTGGGACAAATGATGGCCAGAAAGCAAAAGATTGCTGCGGTGCCAA 385
QY 298 AAGGGAATGCCATGAAATCCGGGGCCGATGTGTGCTGCTCTGCAATTGAGGAA 357
Db 386 AAGGATATCCCGCTTCTGTGGGGCCGTGCTTGGCACTGTGACAGAGGCAAGGTG 445
QY 358 ATGAAGTTGAAAAACCCCGAAGATACAGATCATGAAAGAGGCAAGAGTATCT 417
Db 446 AGTTACACAGAAACCTTGAAAGTTGA---GAGCTGACATGTCCCTGGGACCCC 502
QY 418 GAGCAGATCAGCGCATGACCGGACGTAAAGCGGACATTAAGAAATATATCTTC 477
Db 503 AAGTGTCTGAGGTGATGAGGTGACCTGACACCGGACAGTCCATTCCATGAAATGTT 562
QY 478 AAGGATCGATAGGAACCAAGAGCGGGAATTAATCTCAATCTCTGTGCAATAGGAG 537
Db 563 GTGTCCCGGGGGGCGACCGGACAGAGACCTATTCCTGTGTGAAGGCTTACACGCTG 622
QY 538 TATTAACCCGAGAGTGGCTACTGTAGAGGACCTGAGCCATGCGCGCTTGTCTCTC 597
Db 623 TACCGGCGCGAGAGGGCTACTGTGCAAGGCCAGGCGCCCATTTGCCCTGTCTCATG 682
QY 598 TATCTTCTGAGAGAGATCATTTGGGCACTGTGTGACCT 638
Db 683 CATATGCTGTGAGCAAGCCTTCTGTGCTCTGTGATACAT 723

RESULT 11

US-09-774-528-336

/ Sequence 336, Application US/09774528
/ Patent No. 6743619
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Zhou, Ping
/ APPLICANT: Goodrich, Ryle
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Zhang, Jie

```

; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aiding J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743191el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pc_Fl_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
US-09-774-528-336
```

```

Query Match      3.6%; Score 70; DB 3; Length 1823;
Best Local Similarity 49.5%; Pred. No. 8.7e-09;
Matches 209; Conservative 0; Mismatches 210; Indels 3; Gaps 1;
```

```

QY 220 CGAAGAGCAAGTGGGTGATATGCTGGAGACTGGGAGAAATACAAAGCAGCAG--A 276
DB 322 CGGAGATGAAGTGGGTGAGATGACCTCGCACTGGGAGAAACATGTCCCGCGGTAC 381
QY 277 AAGCTCATAGATGAGCGGTACAAAGGAAATGCCCATGAATCCGGGGCCCGATGTGTCA 336
DB 382 AAGAAGTTAAGATGACATGTCGCGAAGAGCATCCCGTGTGCTGCGCGCCGATGCTGG 441
QY 337 GTCTCTGAACTTGAAGAAATGAAGTTGAAAAACCCCGAAGATACAGATCATGAAG 396
DB 442 CCCCTGTTGTGTGGGGCCCATGTGTGCGAAGAAACAGCCCTGGCACTTATCAGAGCTG 501
QY 397 GAGAAAGGCAAGAGTATCTGAGCACTCCAGCGCATCCGAGAGTAAAGGGGAGCA 456
DB 502 GCAGAGGCCCTTGAGAACCCAGATGATGAGAACCATTTGCAAGGACCTGCAACCTCAA 561
QY 457 TTAAGAAGCATATATTCTTCAGGGATGATACGGAACCAAGCAGCGGAACTACTCCAC 516
DB 562 TTCCTCTGACAGAGATTTTGTGTGCTCAGAGGCCAGCGGACAGAGGGCTCTGTGAG 621
QY 517 ATCTCTGCGATATGAGAGTATTAACCCGAGGTGGCTTACTGCAAGGACCTGAGCCAC 576
DB 622 GTGCTCAAGGCGCTACACCCCTGTATCGACCGAGCAGGGCTACTGCGAAGGCC 681
QY 577 ATGCGCGCTTGTCTCTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 636
DB 682 GTGCTGTGTGTGTCTCATATGACCTGCCCCCAGAGAGAGGCTTCTGTGTGTGTGTG 741
QY 637 CT 638
DB 742 AT 743
```

```

RESULT 12
US-10-120-988-336
; Sequence 336, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyun
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6919193el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
```

```

; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pc_Fl_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
US-10-120-988-336
```

```

Query Match      3.6%; Score 70; DB 3; Length 1823;
Best Local Similarity 49.5%; Pred. No. 8.7e-09;
Matches 209; Conservative 0; Mismatches 210; Indels 3; Gaps 1;
```

```

QY 220 CGAAGAGCAAGTGGGTGATATGCTGGAGACTGGGAGAAATACAAAGCAGCAG--A 276
DB 322 CGGAGATGAAGTGGGTGAGATGACCTCGCACTGGGAGAAACATGTCCCGCGGTAC 381
QY 277 AAGCTCATAGATGAGCGGTACAAAGGAAATGCCCATGAATCCGGGGCCCGATGTGTCA 336
DB 382 AAGAAGTTAAGATGACATGTCGCGAAGAGCATCCCGTGTGCTGCGCGCCGATGCTGG 441
QY 337 GTCTCTGAACTTGAAGAAATGAAGTTGAAAAACCCCGAAGATACAGATCATGAAG 396
DB 442 CCCCTGTTGTGTGGGGCCCATGTGTGCGAAGAAACAGCCCTGGCACTTATCAGAGCTG 501
QY 397 GAGAAAGGCAAGAGTATCTGAGCACTCCAGCGCATCCGAGAGTAAAGGGGAGCA 456
DB 502 GCAGAGGCCCTTGAGAACCCAGATGATGAGAACCATTTGCAAGGACCTGCAACCTCAA 561
QY 457 TTAAGAAGCATATATTCTTCAGGGATGATACGGAACCAAGCAGCGGAACTACTCCAC 516
DB 562 TTCCTCTGACAGAGATTTTGTGTGCTCAGAGGCCAGCGGACAGAGGGCTCTGTGAG 621
QY 517 ATCTCTGCGATATGAGAGTATTAACCCGAGGTGGCTTACTGCAAGGACCTGAGCCAC 576
DB 622 GTGCTCAAGGCGCTACACCCCTGTATCGACCGAGCAGGGCTACTGCGAAGGCC 681
QY 577 ATGCGCGCTTGTCTCTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 636
DB 682 GTGCTGTGTGTGTCTCATATGACCTGCCCCCAGAGAGAGGCTTCTGTGTGTGTGTG 741
QY 637 CT 638
DB 742 AT 743
```

```

RESULT 13
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: Patent In Release #1.0, Version #1.25
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/08/232,463
5 FILING DATE:
6 CLASSIFICATION: 435
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US/07/935,313
9 FILING DATE:
10 APPLICATION NUMBER: EP 91 114 300.6
11 FILING DATE: 26-AUG-1991
12 ATTORNEY/AGENT INFORMATION:
13 NAME: BENT, Stephen A.
14 REGISTRATION NUMBER: 29,768
15 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: (703) 936-9300
18 TELEFAX: (703) 683-4109
19 TELEX: 899149
20 INFORMATION FOR SEQ ID NO: 14:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 7218 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: single
25 TOPOLOGY: linear
26 IMMEDIATE SOURCE:
27 CLONE: pTZgpt-F18
28 US-08-232-463-14

```

[illegible]

```

; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9899
;
; LENGTH: 806
;
; TYPE: DNA
;
; ORGANISM: Drosophila melanogaster
US-09-270-767-9899

```

	Query Match	3.2%	Score 63.4	DB 3	Length 806
	Best Local Similarity	55.1%	Pred. No. 3,98-07		
	Matches 124	Conservative 0	Mismatches 101	Indels 0	Gaps 0
Qy	433 ATGCAGCCGGGACGTAAAGCGGACATTAAAGAGCATATATCTTCAGGAGATGATACGA	492			
Db	668 ATTCAGGCGGGACATCCATCGCATCTTCCCGCGGACAAATCTTCAAGAAATTGGCCGT	609			
Qy	493 ACCAAGCAGGCGGGAACTACTCCATCATCTCCTGGCATATGAGATATTAACCCGGAGGTG	552			
Db	608 TCTGGCCAAAGTGCACCTTTTAAAGTGTCCAAAGCGTATGCCGTTTCATGACAGGAGTT	549			
Qy	553 GGCTACTGCAAGGACCTGAGCCACATCGCGCCTGTTCCTCTTATCTTCTTGAGAG	612			
Db	548 GGATATTGTCAAGGCTCTAAGTTTCATAGCAGCTATGTCTCTTTCATATGCTTGAGAG	489			
Qy	613 GATGATCTTCTGGGACATCGTGCACACTGCTGCGCCAGTAGAAGGAC	657			
Db	488 GATCCGTTCTGTGTCTGAGAGCCCTTATGTACACCTATGGAGCTC	444			

```

RESULT 15
US-08-363-300-1
; Sequence 1, Application US/08363300
; Patent No. 5709927
; GENERAL INFORMATION:
; APPLICANT: Zon, Leonard and Richardson, Paul
; TITLE OF INVENTION: Tbc1 Gene and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,300
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04590/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4039 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..3829
;
;
;
Query Match          2.7%; Score 53.6; DB 2; Length 4039;
Best Local Similarity 49.6%; Prid. No. 0.00039;

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:32:54 ; Search time 5318 Seconds
(without alignments)
17279.020 Million cell updates/sec

Title: US-10-071-838-1
Perfect score: 1964
Sequence: 1 atggagcgtgtagagtcgcgc.....aaatgttcctgtgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1948	99.2	2072	4	HSMD01828
2	1585.8	80.7	2497	4	BC047739
3	859.8	43.8	986	5	BX346181
4	841.2	42.8	915	5	BQ923700
5	814.6	41.5	909	3	BI753688
6	797.6	40.6	954	5	BUS26609
7	797.4	40.6	919	3	BI522600
8	783.6	39.9	868	6	CD653466
9	751.6	38.3	881	5	BUS42901
10	741.2	37.7	818	3	BI597531
11	732.6	37.2	930	5	BQ920917
12	730.2	36.9	1013	3	BM560320
13	724.8	36.6	839	3	BI601698
14	718.2	36.3	1141	3	BM553146
15	713.6	36.3	831	3	BI828272
16	707.4	36.0	709	2	BI088323
17	701.8	35.7	1079	3	BM563571
18	700.8	35.7	732	6	CA748642
19	696.6	35.5	727	1	AL042933
20	694.4	35.4	727	1	AL042933
21	686.6	35.0	875	5	AQ788033
22	675	34.4	741	5	BU688833

23	663.8	33.8	667	3	BI827136
24	651.4	33.2	890	6	CA487310
25	645.8	32.9	849	3	BI829764
26	641	32.6	667	6	CB241156
27	640.4	32.6	882	3	BI520522
28	635.2	32.3	680	2	BE676151
29	632.4	32.2	634	7	CN262008
30	632	32.2	964	4	BM914524
31	628.4	32.0	630	3	BM842553
32	627.4	31.9	629	3	BM843014
33	626.8	31.9	669	1	AI867050
34	626.4	31.9	628	3	BM762465
35	625.4	31.8	654	5	BU625940
36	622.4	31.7	742	3	BI832051
37	620	31.6	628	8	DN999748
38	610.2	31.1	633	8	DR067725
39	609.6	31.0	616	9	AQ420821
40	609.4	31.0	700	1	AI555979
41	604.4	30.8	655	5	BU607318
42	603.4	30.7	656	3	BQ007813
43	600	30.5	637	1	AM440435
44	595	30.2	802	7	CR766757
45	591.8	30.1	650	3	BI830812

ALIGNMENTS

RESULT 1	HSMD01828	2072 bp	mRNA	linear	HTC 22-SEP-2004
LOCUS	HSMD01828				
DEFINITION	Homo sapiens mRNA; cDNA DKFZp434P2235 (from clone DKFZp434P2235).				
ACCESSION	AL136860				
VERSION	AL136860.1	GI:12053224			
KEYWORDS	HTC.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
REFERENCE	1 (bases 1 to 2072)				
AUTHORS	Pouscka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weill,B., Amd,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.				
CONSTRM	The German cDNA Consortium				
TITLE	Direct Submission				
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp434P2235) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434P2235 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ .				
FEATURES	Location/Qualifiers				
source	1..2072				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="RZPD:DKFZp434P2235"				
	/db_xref="taxon:9606"				
	/clone="DKFZp434P2235"				
	/tissue_type="testis"				
	/clone_lib="434 (synonym: htes3). Vector pSport1, host DH10B, sites NotI + SalI"				
	/dev_stage="adult"				
gene	1..2072				
	/note="TBC1 domain family, member 3"				
CDS	/gene="DKFZp434P2235"				
	99..1748				

```
/gene="DKFZp434P2235"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66794.1"
/db_xref="GI:12053225"
/db_xref="UniProt/Swiss-Prot:O81ZP1"
/translation="MDVVEVAGSWMAQERBDIIMKYEKGHRAGLIPEDGKPKPRSYNN
NDHIGLVHETELPPLTAREAKQIRREISRKSKVDMLGDEKYSRKLIDRAYKM
PMNIRGPMWVILNTEEMKLNKPGROYIMKEGKXSEHIIORIDRVSGTLRKHI
DRYGTRORELIHLLAYEYKPYECRODLSHIALPLVYIPEEDARFALVOLIASER
HSLGFEHSPNGSTVQGLDQDEHVATISQPTMGHODKIDCGCCSLICIRLITIG
ISLGLTLRLMDVYLVEGEQALMPTIRIAFKVQRLTKTSKCGWMAFCNRPVDTMAR
DEDTVLKHLRASMKLTTRKQKDLPPPAKPEQSSASPPVPSRSGKTLKGDROAPG
PPAPFPPIWASBPBARBSSTPCPGAVREDTYPVGTQVPSBALAQGGQSWRFL
QNMNMRPLPTDLDEYGPMPHRYDPFROSCWPAISQEDLAPCMQAEHBAERVSAPFA
PSIDSQGTTPRANDBQCAPTSRPLCGLHLESQRPGR"
```

Query Match 99.2%; Score 1948; DB 4; Length 2072;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

DB	Query	Match
Db	1	ATGACGCTGTAGAGTTCGCGGCAATTGTGGGCAACAAGCCAGAGCATCATATG 60
Db	99	ATGACGCTGTAGAGTTCGCGGCAATTGTGGGCAACAAGCCAGAGCATCATATG 158
Db	61	AAATACGAAAGGACACCGAGCTGGGCTGCAGAGGACAGAGGCGCTTAAGCTTTTGA 120
Db	159	AAATACGAAAGGACACCGAGCTGGGCTGCAGAGGACAGAGGCGCTTAAGCTTTTGA 218
Db	121	AGCTACACACACACGTCGATCATTTGGGATGTGACATGACGAGCTGCTCTG 180
Db	219	AGCTACACACACACGTCGATCATTTGGGATGTGACATGACGAGCTGCTCTG 278
Db	181	ACTGCGGCGGAGCGGAGCAATTCGCGGAGATCAGCCGAAAGACAGTGGTGAT 240
Db	279	ACTGCGGCGGAGCGGAGCAATTCGCGGAGATCAGCCGAAAGACAGTGGTGAT 338
Db	241	ATGCTGGAGAGCTGGGGAATTAACAAAGCAGGAAAGCTCATAGATCGAGCTACAG 300
Db	339	ATGCTGGAGAGCTGGGGAATTAACAAAGCAGGAAAGCTCATAGATCGAGCTACAG 398
Db	301	GGAATGCCATGAAATCCTCGGGGCGCGATGTGTCACTCTCTGAACTTTGAGAAATG 360
Db	399	GGAATGCCATGAAATCCTCGGGGCGCGATGTGTCACTCTCTGAACTTTGAGAAATG 458
Db	361	AAAGTTGAAAAACCCCGGAAGATACCAATTCATGAAGAGAGGAGAGGTCTCTGAG 420
Db	459	AAAGTTGAAAAACCCCGGAAGATACCAATTCATGAAGAGAGGAGAGGTCTCTGAG 518
Db	421	CACATCGAGGCGATCGACCGGAGCGTAAGCGGAGCATTTAAGGAGCATATTCTTCAAG 480
Db	519	CACATCGAGGCGATCGACCGGAGCGTAAGCGGAGCATTTAAGGAGCATATTCTTCAAG 578
Db	481	GATGATACGAGAACCAAGCAGCGGAACTTACTCAATCTCTGAGCATATGAGAGAT 540
Db	579	GATGATACGAGAACCAAGCAGCGGAACTTACTCAATCTCTGAGCATATGAGAGAT 638
Db	541	AAACCGAGGTGGCTACTGAGGAGCTTGAAGCAATCGCGCTTGTCTCTCTAT 600
Db	639	AAACCGAGGTGGCTACTGAGGAGCTTGAAGCAATCGCGCTTGTCTCTCTAT 698
Db	601	CTTCTGAGAGAGATGATCTGGGCACTGGTGAAGCTGAGGCAAGAGGACATCC 660
Db	699	CTTCTGAGAGAGATGATCTGGGCACTGGTGAAGCTGAGGCAAGAGGACATCC 758
Db	661	CTGACGAGATTTCACAGCCAAATGCGCGGACCGTCAAGGAGCTCAAGACCAAGAG 720
Db	759	CTGACGAGATTTCACAGCCAAATGCGCGGACCGTCAAGGAGCTCAAGACCAAGAG 818
Db	721	CATGTGTAGCAGCTCAACCAAGACCATGGGAGCTCAGACCAAGAAAGATCTATGT 780
Db	819	CATGTGTAGCAGCTCAACCAAGACCATGGGAGCTCAGACCAAGAAAGATCTATGT 878

Db	878	GAGCAGGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATGACGGATCTCTCGGG 840
Db	879	GAGCAGGTTCCTCCGTTAGGCTGCTCATCCGATATTTGATGACGGATCTCTCGGG 938
Db	841	CTACCTCTGCGCTGTGGGAGCTGTATCTGTATGAAAGGCAACAGCGCTGATGCGGATA 900
Db	939	CTACCTCTGCGCTGTGGGAGCTGTATCTGTATGAAAGGCAACAGCGCTGATGCGGATA 998
Db	901	ACAAGAAATGCGCTTTAAGTTACGAGAACCGCTCAAGAAAGCTCAGATGAGCGCG 960
Db	999	ACAAGAAATGCGCTTTAAGTTACGAGAACCGCTCAAGAAAGCTCAGATGAGCGCG 1058
Db	961	TGGGCACTTTTTCGAAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC 1020
Db	1059	TGGGCACTTTTTCGAAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC 1118
Db	1021	AAGCATCTTAGGCGCTCTATGAAAGAACTTAACAAAGAGGAGGAGCTTGCCACCCCA 1080
Db	1119	AAGCATCTTAGGCGCTCTATGAAAGAACTTAACAAAGAGGAGGAGCTTGCCACCCCA 1178
Db	1081	GCCAAACCCGAGCAAGGAGTGTGCGCATCCAGGCTGTGCGGCTTCAAGTGGCGGAG 1140
Db	1179	GCCAAACCCGAGCAAGGAGTGTGCGCATCCAGGCTGTGCGGCTTCAAGTGGCGGAG 1238
Db	1141	ACCCTCTGCAAGGAGGAGCAGGACAGGCGCTCCAGGCGCAACAGCCGAGTTCCCGGCGCC 1200
Db	1239	ACCCTCTGCAAGGAGGAGCAGGACAGGCGCTCCAGGCGCAACAGCCGAGTTCCCGGCGCC 1298
Db	1201	ATTTGTGACCTTCCCGGCAAGGAGCACTGTCTTCCACACCTGTCTGTGGGAGCT 1260
Db	1299	ATTTGTGACCTTCCCGGCAAGGAGCACTGTCTTCCACACCTGTCTGTGGGAGCT 1358
Db	1261	GTCGCGGAGAAACACTAAGCTGTGGGAGCTCAGGAGTGTGCGGCGCGGCTGGGCTCAG 1320
Db	1359	GTCGCGGAGAAACACTAAGCTGTGGGAGCTCAGGAGTGTGCGGCGCGGCTGGGCTCAG 1418
Db	1321	GGAGAACCTTCAAGGAGTCTGGAATTCCTGCAATGGAATCTCAATGCCCGCTCCCAAG 1380
Db	1419	GGAGAACCTTCAAGGAGTCTGGAATTCCTGCAATGGAATCTCAATGCCCGCTCCCAAG 1478
Db	1381	GACCTGAGCGTAAGAGGCGCTTGTGCTGCGCATTAATGATTTCAACAGAGCTGTGGGCT 1440
Db	1479	GACCTGAGCGTAAGAGGCGCTTGTGCTGCGCATTAATGATTTCAACAGAGCTGTGGGCT 1538
Db	1441	CGTGCAATTCGCAAGAGGAGCAGCTGCGGCGGCTGCGGAGGCTGAACAACCGTGCGGAG 1500
Db	1539	CGTGCAATTCGCAAGAGGAGCAGCTGCGGCGGCTGCGGAGGCTGAACAACCGTGCGGAG 1598
Db	1501	CGGATGAGATCGGCTTTCGCTGCAACGAGCATGATTTCCGACAGGAGCAACCCCTTCAGA 1560
Db	1599	CGGATGAGATCGGCTTTCGCTGCAACGAGCATGATTTCCGACAGGAGCAACCCCTTCAGA 1658
Db	1561	GCTAGGAGCAACAGCGCTGTGTCTCCACTCAGGCGCTTGTCTGTGCGGCTTCACTTG 1620
Db	1659	GCTAGGAGCAACAGCGCTGTGTCTCCACTCAGGCGCTTGTCTGTGCGGCTTCACTTG 1718
Db	1621	GAAAGTTCTCAAGTTCCCTCAGGCTTCTGAACCATCTGGGCGAGGCTCATAGCTGGAT 1680
Db	1719	GAAAGTTCTCAAGTTCCCTCAGGCTTCTGAACCATCTGGGCGAGGCTCATAGCTGGAT 1778
Db	1681	AATTTCCCTGAGCTTAAACCAACCAAGCAAGCTTCCGCTCTCTTTATTTTGTGTTAA 1740
Db	1779	AATTTCCCTGAGCTTAAACCAACCAAGCAAGCTTCCGCTCTCTTTATTTTGTGTTAA 1838
Db	1741	CTTATGAAATGTATTAAGAAAGTGCAGCTCGAAGAGATTCAGAGATGAAACACACC 1800
Db	1839	CTTATGAAATGTATTAAGAAAGTGCAGCTCGAAGAGATTCAGAGATGAAACACACC 1898
Db	1801	AGACCCCAAGTTCACAAAGCCAAACCATGCGGCGCTCCAGGACCCCGAGCCCAAGAC 1860
Db	1899	AGACCCCAAGTTCACAAAGCCAAACCATGCGGCGCTCCAGGACCCCGAGCCCAAGAC 1958

QY 1861 CATCTCTGAAATTCGACGACACCGTGACCTGCTTGTACTTAACTGATGAAG 1920
| | | | |
Db 1959 CATCGTTCTGAATTCGACGACACCGTGACCTGCTTGTACTTAACTGATGAAG 2018
| | | | |
QY 1921 ATACTACTCTTCACTGTTTGAATTAATGTTTCTCTGTTGAATG 1964
| | | | |
Db 2019 ATACACCACTTCATGTTTGAATTAATGTTTCTCTGTTGAATG 2062
| | | | |
RESULT 2
BC047739 2497 bp mRNA linear HTC 20-MAR-2003
LOCUS Homo sapiens, similar to Rab GTPase-activating protein PRC17, clone
DEFINITION IMAGE:5743752, mRNA.
ACCESSION BC047739
VERSION BC047739.1 GI:29126829
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 2497)
Strausberg, R.
Direct Submission
Submitted (03-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAX Plate: 98 Row: d Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984
This clone has the following problem: retained intron.
Location/Qualifiers
1.2497
Source
1. /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5743752"
/tissue_type="Brain, adult medulla"
/clone_id="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Query Match 80.7%; Score 1585.8; DB 4; Length 2497;
Beet Local Similarity 89.8%; Pred. No. 0;
Matches 1796; Conservative 0; Mismatches 12; Indels 193; Gaps 1;
QY 157 CATGAGACGAGCTGCTCTCTCTGACTGCGCGGAGCGCAAGCAATTGCGCGGAGATC 216
| | | | |
Db 468 CGTGAAGCGGAGCTGCTCTCTCTGACTGCGCGGAGCGCAAGCAATTGCGCGGAGATC 527
| | | | |
QY 217 ACCCGAAAGACGAGTGGGTGATGATCTGGGAGACTGGGAGAAATACAAAGACGACGAC 276
| | | | |
Db 538 ACCCGAAAGACGAGTGGGTGATGATGCTGGGAGACTGGGAGAAATACAAAGACGACGAC 587
| | | | |
QY 277 AAGCTCATGATGAGCGTACAAAGGAATGCCATCAATCCGGGGCCCGATGTGTGTA 336
| | | | |

Db 588 AAGCTATGATGATGACGCTACAAAGGAATGCCATCAATCCGGGGCCCGATGTGTGTA 647
| | | | |
QY 337 GTCTCTCTGAACATTGAGGAATGAATGTAAGAAACCCCGGAAGATACGATCATGAG 396
| | | | |
Db 648 GTCTCTCTGAACATTGAGGAATGAATGTAAGAAACCCCGGAAGATACGATCATGAG 707
| | | | |
QY 397 GAGAAGGGCAAGGCTCATCTGAGCATCCGCGCATGAGCCGGGACCGGACCGGAGCA 456
| | | | |
Db 708 GAGAAGGGCAAGGCTCATCTGAGCATCCGCGCATGAGCCGGGACCGGAGCGGAGCA 767
| | | | |
QY 457 TTAAGAGCATATATTTCTTCAGGGATCGATACGAAACGAAGGCGGGAACTACTCCAC 516
| | | | |
Db 768 TTAAGAGCATATATTTCTTCAGGGATCGATACGAAACGAAGGCGGGAACTACTCCAC 827
| | | | |
QY 517 ATCTCTCTGGCATATGAGAGATTAACCCGAGAGTGGGCTACTGACAGGACCTGAGCCAC 576
| | | | |
Db 828 ATCTCTCTGGCATATGAGAGATTAACCCGAGAGTGGGCTACTGACAGGACCTGAGCCAC 887
| | | | |
QY 577 ATGCGCGCTGTGTTCTCTCTATCTTCTGAGAGAGATTCATTCGGGACCTGTGACG 636
| | | | |
Db 888 ATGCGCGCTGTGTTCTCTCTATCTTCTGAGAGAGATTCATTCGGGACCTGTGACG 947
| | | | |
QY 637 CTGCTGGCCAGTGAAGGACCTCCCTGACAGGATTTACAGCCCAATGGCGGACCGTC 696
| | | | |
Db 948 CTGCTGGCCAGTGAAGGACCTCCCTGACAGGATTTACAGCCCAATGGCGGACCGTC 1007
| | | | |
QY 697 CAGGGGCTCCAGAACCAAGAGAGCATGTGTAGCCAGCTCACACCCAAACCATGGGG 756
| | | | |
Db 1008 CAGGGGCTCCAGAACCAAGAGAGCATGTGTAGCCAGCTCACACCCAAACCATGGGG 1067
| | | | |
QY 757 CATCAGAACAAAGAAATCTATGTGGGCAAGTGTCCCGGTAGGCTGCTCATCCGATA 816
| | | | |
Db 1068 CATCAGAACAAAGAAATCTATGTGGGCAAGTGTCCCGGTAGGCTGCTCATCCGATA 1127
| | | | |
QY 817 TTGATTGAGGGAATCTCTCTCGGGCTCACCCCTGCGCTGTGGGACGTGATCTGTAGAA 876
| | | | |
Db 1128 TTGATTGAGGGAATCTCTCTCGGGCTCACCCCTGCGCTGTGGGACGTGATCTGTAGAA 1187
| | | | |
QY 877 GCGGAACAGGCGTGTATGTCGATTAACAGAAATGCTTTAAGTTGACGAGA----- 928
| | | | |
Db 1188 GCGGAACAGGCGTGTATGTCGATTAACAGAAATGCTTTAAGTTGACGAGTAAGTCT 1247
| | | | |
QY 929 ----- 928
| | | | |
Db 1248 ACGTGTGCCAGCGGGGCTGGGGAGCCCTGGGGTCAGACCCGACTGGCCGAGGCGAG 1307
| | | | |
QY 929 ----- 928
| | | | |
Db 1308 CTTCCTCACACTGTCTCATGATCCTCTGTCTGSCCCAGAGGAGGTCTGGCCAGGTGG 1367
| | | | |
QY 929 ----- 928
| | | | |
Db 1368 GCTGGGACAGACACTGTGACACCGAGCCATCCCCACATGACCCAGATGAAGTAGAG 1427
| | | | |
QY 929 -----AGCGCTTCAGAAAGCTCCAGTGTGTGCGCCGTGGGCACTTTTTCACCGGTT 983
| | | | |
Db 1428 GTGTAGCGCTTCAGAAAGCTCCAGTGTGTGCGCCGTGGGCACTTTTTCACCGGTT 1487
| | | | |
QY 984 CGTTGATCTCTGGGCGAGGATGAGGACACTGTGCTCAAGCACTTAAGGCGCTCTTGA 1043
| | | | |
Db 1488 CGTTGATCTCTGGGCGAGGATGAGGACACTGTGCTCAAGCACTTAAGGCGCTCTTGA 1547
| | | | |
QY 1044 GAAACTTAACAAGAAACGAGGGGAGCTGTGCAACCCCGACCAACCCGAGCAAGGGTCTG 1103
| | | | |
Db 1548 GAAACTTAACAAGAAAGAGGGGAGCTGTGCAACCCCGACCAACCCGAGCAAGGGTCTG 1607
| | | | |
QY 1104 GGCATTCAGGCTGTGCGGCTTCACTGTGGCGGAAAGACCTCTGTGCAAGGGGACAGGCA 1163
| | | | |
Db 1608 GGCATTCAGGCTGTGCGGCTTCACTGTGGCGGAAAGACCTCTGTGCAAGGGGACAGGCA 1667
| | | | |
QY 1164 GGCCTCTCAGGCGCAACGACCGCGGTTCCCGGGCCATTTGATCAAGCTTCCCGCAGC 1223
| | | | |

Db 1668 GGGCCCTCCAGAGCCCAAGCCGGTTCGCCGCGGCCCAATTTGTCAAGCTTCCCTCCAGC 1727
Qy 1224 GGGACCTGTTCTTCCACACCCTGTCCTGGTGGGGGCTGTCCGGGAAGACCTACCTCCGT 1283
Db 1728 GGGACCTGTTCTTCCACACCCTGTCCTGGTGGGGGCTGTCCGGGAAGACCTACCTCCGT 1787
Qy 1284 GGGCACTCAGGGGTGCGCCAGCCCGGCTTGGCTCAGGGAAGACCTCAGGGTTCTGAG 1343
Db 1788 GGGCACTCAGGGGTGCGCCAGCCCGGCTTGGCTCAGGGAAGACCTCAGGGTTCTGAG 1847
Qy 1344 ATTCTCTGAGTGAACCTCCATGCCCCCTCCCAAGGACCTGAGAGCTAGAGGCCCTTG 1403
Db 1848 ATTCTCTGAGTGAACCTCCATGCCCCCTCCCAAGGACCTGAGAGCTAGAGGCCCTTG 1907
Qy 1404 GTTCCCGCATTTATGATTTCCAGACAGAGCTGCGGTCCGTCATATCCAGAGAGCA 1463
Db 1908 GTTCCCGCATTTATGATTTCCAGACAGAGCTGCGGTCCGTCATATCCAGAGAGCA 1967
Qy 1464 GCTGGCCCTGCTGCGAGGCTGAACAACCTGCGAGAGCGGGTGAATCGGCTTTCGCTGC 1523
Db 1968 GCTGGCCCTGCTGCGAGGCTGAACAACCTGCGAGAGCGGGTGAATCGGCTTTCGCTGC 2027
Qy 1524 ACCAGACAGTGAATTCGACCGAGGCAACCCCTTCAAGCTAGGAGCGAACAGCCGTGTC 1583
Db 2028 ACCAGACAGTGAATTCGACCGAGGCAACCCCTTCAAGCTAGGAGCGAACAGCCGTGTC 2087
Qy 1584 TCCCACTCAAGGGCTTGGCTGCTGCGGCTCCATCTGGAAGATTCTGAGTTCCCTCCAGG 1643
Db 2088 TCCCACTCAAGGGCTTGGCTGCTGCGGCTCCATCTGGAAGATTCTGAGTTCCCTCCAGG 2147
Qy 1644 CTCTAGAACATCTGGGCGAGGCTCAGGCTGATTAATTTCCCTAGGCTTAAACAACC 1703
Db 2148 CTCTAGAACATCTGGGCGAGGCTCAGGCTGATTAATTTCCCTAGGCTTAAACAACC 2207
Qy 1704 AAGCAAGCTTGGCTCTCTCTTTATTTTGGTTAACTTATGAATAATGATTAAGAAAG 1763
Db 2208 AAGCAAGCTTGGCTCTCTCTTTATTTTGGTTAACTTATGAATAATGATTAAGAAAG 2267
Qy 1764 AGTGCACCTGAGAGAGATTGAGATGAGAAACACACAGAGCCCAAGATCAAAAGCCAAC 1823
Db 2268 AGTGCACCTGAGAGAGATTGAGATGAGAAACACACAGAGCCCAAGATCAAAAGCCAAC 2327
Qy 1824 CATGCCAGGCGCTCCAGACACCCCAAGCCCAAGCAATGTTCTGAAATTTGACGACA 1883
Db 2328 CATGCCAGGCGCTCCAGACACCCCAAGCCCAAGCAATGTTCTGAAATTTGACGACA 2387
Qy 1884 CCGTGAAGCTGCTTTGTAATCTTAATCATGAGAGATTAATCACTTCAAGTTTGA 1943
Db 2388 CCGTGAAGCTGCTTTGTAATCTTAATCATGAGAGATTAATCACTTCAAGTTTGA 2447
Qy 1944 TAAATGTTCTCTGTTGAATG 1964
Db 2448 TAAATGTTCTCTGTTGAATG 2468

RESULT 3
BX346181/c 986 bp mRNA linear EST 08-Apr-2004
LOCUS BX346181 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0D010YK20 3-PRIME, mRNA sequence.
ACCESSION BX346181
VERSION BX346181.1 GI:30369074
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1. (bases 1 to 986)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Creneau, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8800.x
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0D010B10NP1&c=8800.r.
FEATURES
source
Location/Qualifiers
1. 986
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D010YK20"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_id="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 43.8%; Score 859.8; DB 5; Length 986;
Best Local Similarity 95.4%; Pred. No. 4.9e-219; Indels 7; Gaps 5;
Matches 925; Conservative 9; Mismatches 29; Indels 7; Gaps 5;
Qy 971 TTTCACACCGTTCGTGATATCTGGGCGCAGGATGAGACAGTGTCTCAACATCTTA 1030
Db 979 KTTTAAACCGTTCGTGATATCTGGGCGCAGGATGAGACAGTGTCTCAACATCTTA 921
Qy 1031 GGGCTCTATGAAGAACTTAACAAGAGAGGGGAGCTTCCACCCCGCAACCCG 1090
Db 920 GGCC--TCTAKARAAACTTAACAAGAGAGGGGAGCC--KCAACCCCGCAACCCG 865
Qy 1091 AGAAGGGGTGTGGGACATCCAGGCTGTGCGGCTTCAAGTGGGGAAGACCTCTGCA 1150
Db 864 AGAAGGGGTGTGGGACATCCAGGCTGTGCGGCTTCAAGTGGGGAAGACCTCTGCA 805
Qy 1151 AGGGGAGCAGGAGGCGCTTCAGGCGCCACAGCCCGGTTCCGCGGCCATTTGTGAC 1210
Db 804 AGGGGAGCAGGAGGCGCTTCAGGCGCCACAGCCCGGTTCCGCGGCCATTTGTGAC 745
Qy 1211 CTTCGCCGCAAGGAGCCTGCTTTCACACCTGTCTCTGTGGGGCTGTCCGGAG 1270
Db 744 CTTCGCCGCAAGGAGCCTGCTTTCACACCTGTCTCTGTGGGGCTGTCCGGAG 685
Qy 1271 ACACTTACCTGTGGGACATCAGGGTGTGCGCAGCCCGGCTTGGCTCAGGAGGACCTC 1330
Db 684 ACACTTACCTGTGGGACATCAGGGTGTGCGCAGCCCGGCTTGGCTCAGGAGGACCTC 625
Qy 1331 AGGGTTCCTGAGATTCCTGCAAGTGAATCAATGCCCCCTCCCAAGCAGACTGAGC 1390
Db 624 AGGGTTCCTGAGATTCCTGCAAGTGAATCAATGCCCCCTCCCAAGCAGACTGAGC 565
Qy 1391 TAGAGGCGCTTGGTTCCGCAATTAATGATTTCAAGACAGAGCTGTGGTCCGTCAT 1450
Db 564 TAGAGGCGCTTGGTTCCGCAATTAATGATTTCAAGACAGAGCTGTGGTCCGTCAT 505
Qy 1451 CCCAGAGAGCAGAGTGGCCCGCTGCGAGAGCTGAACCCCTGCGAGCGGGTGAAT 1510
Db 504 CCCAGAGAGCAGAGTGGCCCGCTGCGAGAGCTGAACCCCTGCGAGCGGGTGAAT 445
Qy 1511 CGGCTTTCGTGACCCAGACATGATTCGACAGGAGCAACCCCTTCAAGACTAGGAGC 1570
Db 444 CGGCTTTCGTGACCCAGACATGATTCGACAGGAGCAACCCCTTCAAGACTAGGAGC 385
Qy 1571 AACAGCGTGTGCTCCCACTCAGGCGCTTGGCTTGGGCTTCACCTTGAAGATTCTC 1630

Db 384 AACAGCCGTGCTCCCACTCAGGGCTTGCTGCGGCTCCACTTGAAGATTCTC 325
 QY 1631 AGTTCCCTCAGGCTTCTAGAGCATCTGGCCAGGGCTCATGGCTGATATTTCCCTA 1690
 Db 324 AGTTCCCTCAGGCTTCTAGAGCATCTGGCCAGGGCTCATGGCTGATATTTCCCTA 265
 QY 1691 GGCTTAACAACCCAGCAGCTTCCGCTCTCTTTTATTTTGTAACTTATGAAA 1750
 Db 264 GGCTTAACAACCCAGCAGG- TTGGATCTCTGTTTATTTTGTAACTTATGAAA 206
 QY 1751 TGTATTAGAAAGATGACCTGAGAGATTCAGAGATGAACACACACCCAGCA 1810
 Db 205 TGTATTAGAAAGATGACCTGAGAGATTCAGAGATGAACACACACCCAGCA 146
 QY 1811 TCACAAAGCCAGCATGCGCCGCTCCAGACACCCCGACCCAGCATGCTCTG 1870
 Db 145 TCACAAAGCCAGCATGCGCCGCTCCAGACACCCCGACCCAGCATGCTCTG 86
 QY 1871 AATTCTGACGACACCGTGAAGCTGCTGCTTTGTTAACTC-ATGGAAGATTAAC 1929
 Db 85 AATTCTGACGACACCGTGAAGCTGCTGCTTTGTTAACTC-ATGGAAGATTAAC 26
 QY 1930 TTCACGTTT 1939
 Db 25 CTCNATGTT 16

RESULT 4
 BQ923700 915 bp mRNA linear EST 20-AUG-2002
 LOCUS AGENCOURT 8798484 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6425622
 DEFINITION 5', mRNA Sequence.

ACCESSION BQ923700
 VERSION BQ923700.1 GI:22338731
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 915)

REFERENCE
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@ds-r@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2607 row: 1 column: 07
 High quality sequence stop: 639.
 Location/Qualifiers

FEATURES

source

1..915
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6425622"
 /rname_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_101"
 /note="Organ: lung; Vector: pOT87; Site 1: EcorI; Site 2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned
 into EcorI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match 42.8%; Score 841.2; DB 5; Length 915;
 Best Local Similarity 98.5%; Pred. No. 4,7e-214;
 Matches 890; Conservative 0; Mismatches 10; Indels 4; Gaps 4;

QY 409 AGGTCACTGAGCATCCAGCGCATGACCCGGAAGTAAAGCGGACATTAAAGAAACAT 468
 Db 1 AAGGCATCTGAGCATCCAGCGCATGACCCGGAAGTAAAGCGGACATTAAAGAAACAT 60
 QY 469 ATATTCTTACAGGATGATATCGAACAAGACGCGGAACTATCCACATCTCTCGGCA 528
 Db 61 ATATTCTTACAGGATGATATCGAACAAGACGCGGAACTATCCACATCTCTCGGCA 120
 QY 529 TATGAGAGTATTAACCCGAGGCTGCTCTGAGGGAAGCTGAGGCACTGCGGCTTG 588
 Db 121 TATGAGAGTATTAACCCGAGGCTGCTCTGAGGGAAGCTGAGGCACTGCGGCTTG 180
 QY 589 TTCTCTCTATCTTCTGAGAGATGCAATCTGAGGCACTGAGTGCCTGCGCAGT 648
 Db 181 TTCTCTCTATCTTCTGAGAGATGCAATCTGAGGCACTGAGTGCCTGCGCAGT 240
 QY 649 GAGAGCACTCCCTGAGGGAATTTCAAGCCCAATGGGGGACCGTCCAGGGGCTCCA 708
 Db 241 GAGAGCACTCCCTGAGGGAATTTCAAGCCCAATGGGGGACCGTCCAGGGGCTCCA 300
 QY 709 GACCAACAGAGATGAGGAGCAGTCAACACCCAGACATGGGGGATCAGACAAG 768
 Db 301 GACCAACAGAGATGAGGAGCAGTCAACACCCAGACATGGGGGATCAGACAAG 360
 QY 769 AAGATCTATGAGGAGTGTCCCGGTTAGGCTGCTCATCCGATATTGATTGACGG 828
 Db 361 AAGATCTATGAGGAGTGTCCCGGTTAGGCTGCTCATCCGATATTGATTGACGG 420
 QY 829 ATCTCTCTGAGGCTCACCTCGGCTGAGGAGCTATCTGTTAAGGCCAACAGGG 888
 Db 421 ATCTCTCTGAGGCTCACCTCGGCTGAGGAGCTATCTGTTAAGGCCAACAGGG 480
 QY 889 TTGATGCGATTAACAAATGCGCTTTAAGGTTACAGAGAGGCTCAAGAAAGCTCC 948
 Db 481 TTGATGCGATTAACAAATGCGCTTTAAGGTTACAGAGAGGCTCAAGAAAGCTCC 540
 QY 949 AGGTGTGAGCCGTGAGCAGTTTTCGAACCGGTTGCTGATACCTGAGGACAGGATGAG 1008
 Db 541 AGGTGTGAGCCGTGAGCAGTTTTCGAACCGGTTGCTGATACCTGAGGACAGGATGAG 600
 QY 1009 GACACTGTGCTCAAGCATCTTATGAGGCTTATGAACTAACTAAGAAAGCAGGGGAGC 1068
 Db 601 GACACTGTGCTCAAGCATCTTATGAGGCTTATGAACTAAGAAAGCAGGGGAGC 660
 QY 1069 CTGCAACCCCGAGCCAAACCCAGCAAGGAGTGTGCGCATTCAGGCTGCGGCTTCA 1128
 Db 661 CTGCAACCCCGAGCCAAACCCAGCAAGGAGTGTGCGCATTCAGGCTGCGGCTTCA 720
 QY 1129 CGTGGGGAAGACCTCTGCAAGGGGGAACAGGAGGCTTCAGAGGCCACAGCCGG 1188
 Db 721 CGTGGGGAAGACCTCTGCAAGGGGGAACAGGAGGCTTCAGAGGCCACAGCCGG 779
 QY 1189 TTCCGCGGAGCCATTGTCAGCTTCCCGGCAAGGCACTTCGTTCTTCCACACCTG- 1247
 Db 780 TTCCGCGGAGCCATTGTCAGCTTCCCGGCAAGGCACTTCGTTCTTCCACACCTG- 839
 QY 1248 TCTGTGTGGGCTGTTC- GGGAAAGACCTA- CCTGTGGGCACTAAGGTGTGGCCAGC 1305
 Db 840 TCTGTGTGGGCTGTTC- GGGAAAGACCTA- CCTGTGGGCACTAAGGTGTGGCCAGC 899
 QY 1306 CCGG 1309
 Db 900 CCGG 903

RESULT 5
 B1753688 909 bp mRNA linear EST 25-SEP-2001
 LOCUS B1753688

DEFINITION 603023590F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193959 5', mRNA sequence.

ACCESSION B1753688

VERSION B1753688.1 GI:15745266

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 909)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LLM11485 row: d column: 24
High quality sequence stop: 884.
Location/Qualifiers

FEATURES

Source

1..909

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5193959"

/lab_host="DH10B"

/clone_11b="NIH MGC_114"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 41.5%; Score 814.6; DB 3; Length 909;

Best Local Similarity 97.1%; Pred. No. 6.4e-207;

Matches 883; Conservative 0; Mismatches 19; Indels 7; Gaps 5;

164 CGAGCTGCTCTCTGACTGCGCGGAGCGGAAGCAAAATTCGCGGAGATCAGCCGA 223

1 CGAGCTGCTCTCTGACTGCGCGGAGCGGAAGCAAAATTCGCGGAGATCAGCCGA 60

224 -AGAGCAAGTGGTGGATATCTCTGGGAGACTGGGAGAAATTCAGAAAGAGAGAAAGCTC 282

61 GAGAGCAAGTGGTGGATATCTCTGGGAGACTGGGAGAAATTCAGAAAGAGAGAAAGCTC 120

283 ATAGATGAGGCTGCAAGGAAATCCCATGAAATTCGCGGAGCGGATGTGATCAGTCTC 342

121 ATAGATGAGGCTGCAAGGAAATCCCATGAAATTCGCGGAGCGGATGTGATCAGTCTC 180

343 CTGAACATTTGAGAAATGAAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAG 402

181 CTGAACATTTGAGAAATGAAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAG 240

403 GCGAAGAGTCTATCTGAGCAGATCAAGGATCAAGCGGAGCTTAAGCGGAGCACTTAAG 462

241 GCGAAGAGTCTATCTGAGCAGATCAAGGATCAAGCGGAGCTTAAGCGGAGCACTTAAG 300

463 AAGCATATATTTCTCAGGAGATCGATAGGAAACCAAGAGCGGAGAACTTCAATCTC 522

301 AAGCATATATTTCTCAGGAGATCGATAGGAAACCAAGAGCGGAGAACTTCAATCTC 360

523 CTGGCATATGAGAGATTAACCCGAGAGTGGCTTACTGACAGGAGCACTGAGCCATCGCC 582

Db 361 CTGGCATATGAGAGATTAACCCGAGAGTGGCTTACTGACAGGAGCACTGAGCCATCGCC 420

Qy 583 GCGTTGTTCTCTCCCTATCTTCTGAGAGAGATGATCTTGGGCACTGATGAGAGCTCTG 642

Db 421 GCGTTGTTCTCTCTCTATCTTCTGAGAGAGATGATCTTGGGCACTGATGAGAGCTCTG 480

Qy 643 GCCAGTGAAGAGCACTCTCTGAGAGATTTTCAAGCCCAATGCGGAGACCTTCAAGAG 702

Db 481 GCCAGTGAAGAGCACTCTCTGAGAGATTTTCAAGCCCAATGCGGAGACCTTCAAGAG 540

Qy 703 CTCCAGAACCAAGAGAGCATGTGTAGCCACAGTCAACCCAGAACCAATGCGGAGCATCAG 762

Db 541 CTCCAGAACCAAGAGAGCATGTGTAGCCACAGTCAACCCAGAACCAATGCGGAGCATCAG 600

Qy 763 GACAAGAAAGATCATGTGAGGAGTGTTCGCCGTTAGGAGTCCCTCATCCGATATTGAT 822

Db 601 GACAAGAAAGATCATGTGAGGAGTGTTCGCCGTTAGGAGTCCCTCATCCGATATTGAT 660

Qy 823 GACGGAGATCTCTCGGGCTCACCTGCGCTGTGGAGCGTGTATCTGTAGAGGCGAA 882

Db 661 GACGGAGATCTCTCGGGCTCACCTGCGCTGTGGAGCGTGTATCTGTAGAGGCGAA 720

Qy 883 CAGGCGTTGATGCCG-ATTAACAAGATCGCTTTAAGTTTCAAGAGAGCGCTCACGAA 941

Db 721 CAGGCGTTGATGCCGATTAACAAGATCGCTTTAAGTTTCAAGAGAGCGCTCACGAA 780

Qy 942 GAGCTCCA-GATGAGCGCGTGGG--CAGCTTTTGGCAACGGTGGTTGATGATCTGAGG 998

Db 781 GAGCTTACAGATGTGCGGCGCTGAGGCGCACTTTTTCAGAGGATTCGTAGATCTGAGG 840

Qy 999 CAGGCG-ATGAGAGCACTGTGCTCAAGCATCTTAGGGCTCTTAGAGAAATCAACAAGA 1056

Db 841 CAGGCGATTTGAGAGCACTGTGCTCAAGCATCTTAGGGCTCTTAGAGAAATCAACAAGA 900

Qy 1057 AAGCAGGCG 1065

Db 901 AAGCAGGCG 909

RESULT 6

BUS26609

LOCUS 954 bp mRNA linear EST 13-SEP-2002

DEFINITION AGENCOURT 10181753 NIH_MGC_101 Homo sapiens cDNA clone

IMAGE:6536307 5', mRNA sequence.

ACCESSION BUS26609

VERSION BUS26609.1 GI:22837050

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 954)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LLM2697 row: h column: 03
High quality sequence stop: 627.
Location/Qualifiers

FEATURES

Source

1..954

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5536307"
 /issue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_101"
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
 XhoI; cDNA made by oligo-dt priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."

ORIGIN

Query Match	40.6%;	Score 797.6;	DB 5;	Length 954;
Best Local Similarity	98.1%;	Pred. No. 2,4e-202;		
Matches 828;	Conservative 0;	Mismatches 14;	Indels 2;	Gaps 2;
QY	517	ATCCTCTGGCATATGAGATTAACCCGAGATGGCTACTGACAGGACCTGAGCCAC	576	
DB	1	ATCCTCTGGCATATGAGATTAACCCGAGATGGCTACTGACAGGACCTGAGCCAC	60	
QY	577	ATGCGCGCTTGTCTCTCTCTATCTCTGAGAGAGATCATTTGGGCACTGTGCAG	636	
DB	61	ATGCGCGCTTGTCTCTCTCTATCTCTGAGAGAGATCATTTGGGCACTGTGCAG	120	
QY	637	CTGCTGGCCAGTGAAGGCACTCCCTGACAGGATTTCAACGCCCAATGCGGGACCTGC	696	
DB	121	CTGCTGGCCAGTGAAGGCACTCCCTGACAGGATTTCAACGCCCAATGCGGGACCTGC	180	
QY	697	CAGGGGCTCCAAACCAAGAGAGATGTGTGACCACTGACCAACCAAGACATATGGG	756	
DB	181	CAGGGGCTCCAAACCAAGAGAGATGTGTGACCACTGACCAACCAAGACATATGGG	240	
QY	757	CATCAGAGCAAAAGATCTATGTGGGCACTGTCCCGTTAGGCTGCTCATCCGAGTA	816	
DB	241	CATCAGAGCAAAAGATCTATGTGGGCACTGTCCCGTTAGGCTGCTCATCCGAGTA	300	
QY	817	TTGATTGACGGGATCTCTCTGCGGCTCACTTGGCGCTGTGGGACGTGTATCTGTGAA	876	
DB	301	TTGATTGACGGGATCTCTCTGCGGCTCACTTGGCGCTGTGGGACGTGTATCTGTGAA	360	
QY	877	GGGGAACAGGCGTGTATGCGGATTAACGAATGCGCTTAAAGTTACGAGAGCGCTC	936	
DB	361	GGGGAACAGGCGTGTATGCGGATTAACGAATGCGCTTAAAGTTACGAGAGCGCTC	420	
QY	937	ACGAAGACGTCAAGTGTGCGCGGTGGGCACTTTTTCACACCGGTTCTGTATACCTGG	996	
DB	421	ACGAAGACGTCAAGTGTGCGCGGTGGGCACTTTTTCACACCGGTTCTGTATACCTGG	480	
QY	997	GCCAGGATGAGACACTGTGTCACATCTTAAAGGCTCTATGAAGAACTPAACAGA	1056	
DB	481	GCCAGGATGAGACACTGTGTCACATCTTAAAGGCTCTATGAAGAACTPAACAGA	540	
QY	1057	AAGCAGGGGGAACCTGCAACCCCAAGCAACCCGAGAGAGGTCTGTGGCATCAAGGCTT	1116	
DB	541	AAGCAGGGGGAACCTGCAACCCCAAGCAACCCGAGAGAGGTCTGTGGCATCAAGGCTT	600	
QY	1117	GTCGCGGCTTCACTGTGGCGGGAAGACCTCTGCAAGGGGGAACAGGAGCGCTCCAGGC	1176	
DB	601	GTCGCGGCTTCACTGTGGCGGGAAGACCTCTGCAAGGGGGAACAGGAGCGCTCCAGGC	660	
QY	1177	CCACAGCCCGGTTCCCGCGGCACTTGTGACGCTTCCCGCAAGGAGCACTCGTTCT	1236	
DB	661	CCACAGCCCGGTTCCCGCGGCACTTGTGACGCTTCCCGCAAGGAGCACTCGTTCT	720	
QY	1237	TTCACACCTCTGTCTGTGGGCTGTCCGGGAAGACCTAACCCTGTGGCATCAAGGCT	1296	
DB	721	TTCACACCTCTGTCTGTGGGCTGTCCGGGAAGACCTAACCCTGTGGCATCAAGGCT	780	
QY	1297	GTCGCAAGCCGAG-CCCTGTGCTCAGGAGG-ACCTCAGAGGTTCTGTGAGATTCCTGAGT	1354	
DB	781	GTCGCAAGCCGAGCCCTGTGCTCAGGAGG-ACCTCAGAGGTTCTGTGAGATTCCTGAGG	840	

QY 1355 GGAA 1358
 DB 841 GGGA 844

RESULT 7
 B1522600
 LOCUS
 DEFINITION
 B1522600 919 bp mRNA linear EST 29-AUG-2001
 603175612P1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240090 5',
 mRNA sequence.
 VERSION
 B1522600.1 GI:15347392
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.
 1 (bases 1 to 919)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1605 row: g column: 03
 High quality sequence stop: 834.
 Location/Qualifiers
 1. 919
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5240090"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dt primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH_MGC library."

FEATURES

source

ORIGIN

Query Match	40.6%;	Score 797.4;	DB 3;	Length 919;
Best Local Similarity	96.1%;	Pred. No. 2,7e-202;		
Matches 872;	Conservative 0;	Mismatches 26;	Indels 9;	Gaps 5;
QY	895	CCGATTAACAAGATCCCTTTAAGTTACGAGAAACGCTCAAGAAAGCTCAGGTGT	954	
DB	1	CCGATTAACAAGATCCCTTTAAGTTACGAGAAACGCTCAAGAAAGCTCAGGTGT	60	
QY	955	GGCGGTGGGACAGTTTGTGCAACCGGTTCTGTGATACCTGGGCAAGGATGAGCACT	1014	
DB	61	GGCGGTGGGACAGTTTGTGCAACCGGTTCTGTGATACCTGGGCAAGGATGAGCACT	120	
QY	1015	GTCCTAAGATTTTAAAGGCTCTTAAGAAAGAACTAAAGAAAGAGAGGAGGACCTGCCA	1074	
DB	121	GTCCTAAGATTTTAAAGGCTCTTAAGAAAGAACTAAAGAAAGAGAGGAGGACCTGCCA	180	
QY	1075	CCCCAGCCAAACCCAGAGCAAGGATGTGCGGCTCAAGGCTGTGCGGCTTCACTGAGC	1134	
DB	181	CCCCAGCCAAACCCAGAGCAAGGATGTGCGGCTCAAGGCTGTGCGGCTTCACTGAGC	240	

Oy	1135	GGGAAAGACCCCTGTGCAAGGGGGGACAGGAGGCCCCCTCCAGGCGCCACAGAGCCCGATTCCG	1194
Db	241	GGGAAAGACCTCTGTCAAGGGGGACAGGAGGCCCCCTCCAGGCGCCACAGGCCCGATTCCG	300
Oy	1195	CGGCCCATTTGGTCACTTCCCGGACAGGGGACCTGTGTTCCACACCTGTCTGGT	1254
Db	301	CGGCCCATTTGGTCACTTCCCGGACAGGGGACCTGTGTTCCACACCTGTCTGGT	360
Oy	1255	GGGGGTGTCGGGAGACCTAACCTGTGGGCACTCAAGGGTGTGCCAGCCCGGCTGT	1314
Db	361	GGGGGTGTCGGGAGACCTAACCTGTGGGCACTCAAGGGTGTGCCAGCCCGGCTGT	420
Oy	1315	GCTCAGGAGGACCTCAGGGTTCCTGAGATTCTGCAGTGAACTCATGCCGCTC	1374
Db	421	GCTCAGGAGGACCTCAGGGTTCCTGAGATTCTGCAGTGAACTCATGCCGCTC	480
Oy	1375	CCAACTGACCTGGAAGCTGAGGGGCTTGTGTTCCGCAATTATGATTTTCAACAGACTGC	1434
Db	481	CCAACTGACCTGGAAGCTGAGGGGCTTGTGTTCCGCAATTATGATTTTCAACAGACTGC	540
Oy	1435	TGGGTCGTGCATATCCAGGAGGACAGCTGGGCCCTGTGTCAGAGCTGAACACT	1494
Db	541	TGGGTCGTGCATATCCAGGAGGACAGCTGGGCCCTGTGTCAGAGCTGAACACT	600
Oy	1495	GCGAGCGGGGTGAGATCGGCTTTCGCTGCACCAGCACTGATTCGACCAAGGGCA-CCCC	1553
Db	601	GCGAGCGGGGTGAGATCGGCTTTCGCTGCACCAGCACTGATTCGACCAAGGGCACCCC	660
Oy	1554	CTTCAGAGCTAGGGAGCAAGACGCGTGTGCCACTCCAGGGGCTTGCTCTGTGGGCGCT	1613
Db	661	CTTCAGAGCTAGGGAGCAAGACGCGTGTGCCACTCCAGGGGCTTGCTCTGTGGGCGCT	719
Oy	1614	CCACTTGAAGATTCTCACTTCCCTCCAGGCTTCTAGAA----GCATCTGGGCCAGGGC	1668
Db	720	CCACTTGAAGATTCTCACTTCCCTCCAGGCTTCTAGAAAGCATTTCTGGGCCCAAGCT	779
Oy	1669	TCATGGCTGG-ATATATTCCTCAGGCTTAACAACCAAGCAAGCTTC-GCGCTCGATT	1728
Db	780	CCATGGCTGGCATTAATTCCTCAGGCTTAACAACCAAGCAAGCTTCGCAATCTTCGTT	839
Oy	1727	TATTTTGTATTAACCTTAATGAATAATGATTTAAGAAAGATGAGCTCGAGAGATTCAG	1786
Db	840	TATGTATGTTAAACTTAATGAACCTGTCTTAAGAAAGATGCGTTCGAGAGATTCAG	899
Oy	1787	AGATGGA 1793	
Db	900	AGATGGA 906	

RESULT	8
LOCUS	CD653466
DEFINITION	CD653466 868 bp mRNA linear EST 18-JUN-2003
ACCESSION	AGNCOURT_14539069 NIA Human H1 Embryonic Stem Cell cDNA Library (Long) Homo sapiens CDNA clone IMAGE:30423486 5' , mRNA sequence.
VERSION	CD653466
KEYWORDS	CD653466.1 GI:31891804
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 868) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: csapbs-remail.nih.gov Tissue Procurement: Irene Gintis and Mahendra Rao, NIA
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

	CDNA Library Preparation:	Vylan Piao and Minoru Ko
	CNA library Arrayed by:	The I.M.A.G.E. Consortium (LNLN)
	DNA Sequencing Arrayed by:	Agencourt Bioscience Corporation
	Clone distribution:	MGC C
	One distribution information can be found through:	the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
	Plate:	NIDM504 row: j column: 07
	High quality sequence stop:	695.
FEATURES	Location/Qualifiers	
source	1..868	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone_image="IMAGE:30423486"	
	/tissue_type="Embryonic Stem cells"	
	/cell_line="WA01"	
	/lab_host="MDI0B ('T1 phase-resistant')"	
	/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"	
	/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544191]) from WA01 cell line . Undifferentiated human ES cell line WA01/H1 was obtained from Wicell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEf feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDx1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with RNeasy Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544191] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-gGACTAGTCATGATCGCAGCCGCCCTTTTATTATT-3'] from 3.4% of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lona-linker LR-Sal-I, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal-I-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NciI enzymes and cloned into SalI/NciI site of pCMV-Sports plasmid vector. The average insert size is about 3.6kb."	
ORIGIN		
	Query_Match	39.9%; Score 783.6; DB 6; Length 868;
	Best Local Similarity	97.6%; Pred. No. 1.3e-198;
	Matches	803; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
OY	448 AGCGGACATTAAAGAACAATATTTCTTGAGGATCATACGAACCAGACGGGAA	507
DB	1 AAGGTATCAAGAAGAACATATTTCTTGAGGATCATACGAACCAGACGGGAA	60
OY	508 CTACTCCACATCTCTCTGGCATATGAGAGATTAACCCGAGGTGGGTACTATCGAGGAC	567
DB	61 CTACTCCACATCTCTCTGGCATATGAGAGATTAACCCGAGGTGGGTACTATCGAGGAC	120
OY	568 CTGAGCCAATCGCCGCTTGTCTCTCTATCTCTCTGAGGAGGATTCGAGGCA	627
DB	121 CTGAGCCAATCGCCGCTTGTCTCTCTCTATCTCTCTGAGGAGGATTCGAGGCA	180
OY	628 CTGGTAGAGCTGTGGCAGTAGAGGACCTCCCTGAGGAGATTTCAACGCCAATGGC	687
DB	181 CTGGTAGAGCTGTGGCAGTAGAGGACCTCCCTGAGGAGATTTCAACGCCAATGGC	240
OY	688 GGAGACGTCACAGGGCTTCAAAGACCAAGAGACATGTGTAGCCACGTCACCAACCCAG	747


```
Db      241 GGGAGCCCTCCAGGGGCTCCAGACCAAGAGAGCATGTGAGCCACGTCACAAACCAAG 300
Qy      748 ACCATGGGGGATCAGACCAAGAAAGATCTATGTGGGAGGTTCCTCCGGTTAGGCTGCTC 807
Db      301 ACCATGGGGGATCAGACCAAGAAAGATCTATGTGGGAGGTTCCTCCGGTTAGGCTGCTC 360
Qy      808 ATCCGATATTTGATTTGACGGGATCTCTCTGGGCTCAACCTGCGCTGTGGGAGCTGTAT 867
Db      361 ATCCGATATTTGATTTGACGGGATCTCTCTGGGCTCAACCTGCGCTGTGGGAGCTGTAT 420
Qy      868 CTGTGAGAGGCGCAACAGCGGTGTATGCCGATTAACAAGATCGCTTTAAGTTACAGAG 927
Db      421 CTGTGAGAGGCGCAACAGCGGTGTATGCCGATTAACAAGATCGCTTTAAGTTACAGAG 480
Qy      928 AAGCGCTTCAAGAAAGCTCAGGTGTGGCCCGGGGCAAGTTTTCGAAACCGGTTGTT 987
Db      481 AAGCGCTTCAAGAAAGCTCAGGTGTGGCCCGGGGCAAGTTTTCGAAACCGGTTGTT 540
Qy      988 GATACCTGGGCGAGGATGAGACACTGTGCTCAAGCATCTTAGGGGCTCTATGAAGAA 1047
Db      541 GATACCTGGGCGAGGATGAGACACTGTGCTCAAGCATCTTAGGGGCTCTATGAAGAA 600
Qy      1048 CTAAACAAGAAAGAGGGGAGACTGCAACCCCAAGCCCAAGCAAGGGTGTGCGCA 1107
Db      601 CTAAACAAGAAAGAGGGGAGACTGCAACCCCAAGCCCAAGCAAGGGTGTGCGCA 660
Qy      1108 TCCAGGCTGTGCGGCTTACAGTGTGGGGAAGACCTCTGCAAGGGGGAACAGGAGGCC 1167
Db      661 TCCAGGCTGTGCGGCTTACAGTGTGGGGAAGACCTCTGCAAGGGGGAACAGGAGGCC 720
Qy      1168 CTTCAAGGCGGCGGCTTCCGCGGCGGCGCAATTTGATGAGCTT-CGCCGCGAGGGG 1226
Db      721 CTTCAAGGCGGCGGCTTCCGCGGCGGCGCAATTTGATGAGCTT-CGCCGCGAGGGG 780
Qy      1227 ACCTGCTTTTTCACACCTGCTGCTGTGGGCTGTCCGGGAA 1269
Db      781 ACCTGCTTTTTCACACCTGCTGCTGTGGGCTGTCCGGGAA 823
```

```
RESULT 9
BUS42901      881 bp      mRNA      linear      EST 13-SEP-2002
LOCUS      AGENCOURT 10334768 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574864
DEFINITION 5', mRNA sequence.
ACCESSION  BUS42901
VERSION     BUS42901.1 GI:22853384
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-roma@nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2770 row: n column: 16
High quality sequence stop: 642.
Location/Qualifiers

FEATURES

1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:6574864"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site: 1; XhoI; Site: 2; EcoRI; cDNA made by oligo-dr priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      38.3%; Score 751.6; DB 5; Length 881;
Best Local Similarity 95.1%; Pred. No. 5.2e-190;
Matches 808; Conservative 0; Mismatches 39; Indels 3; Gaps 3;

Qy      242 TGGTGGGAGACTGGGAGAAATACAAAAGCAGAAAGCTCATAGATCGAGCGTAAAG 301
Db      2 TGGTGGGAGACTGGGAG-ATATCAAAAAGCAGAAAGCTCATAGATCGAGCGTAAAG 60
Qy      302 GAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTCTGAACTTTAGGAAATGA 361
Db      61 GAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACTTTAGGAAATGA 120
Qy      362 AGTTGAAAAACCCCGAAGATATACAGATCATGAAGAAAGGCGCAGAGGTCACTGAGC 421
Db      121 AGTTGAAAAACCCCGAAGATATCAAGATCATGAAGAAAGGCGCAGAGGTCACTGAGC 180
Qy      422 ACATCCAGGCGATCGACCGGGAGCTTAAGCGGAGCATTTAAGAGAGCATATTTCTTCAAGG 481
Db      181 ACATCCAGGCGATCGACCGGGAGCTTAAGCGGAGCATTTAAGAGAGCATATTTCTTCAAGG 240
Qy      482 ATGATACGGAACCAAGACGCGGAACTACTCCACATCTCTGCGATATGAGAGATATA 541
Db      241 ATGATACGGAACCAAGACGCGGAACTACTCCACATCTCTGCGATATGAGAGATATA 300
Qy      542 ACCCGAGGTGGGCTACTGAGGAGCTGAGGCAATCGCCGCTTCTCTCTATC 601
Db      301 ACCCGAGGTGGGCTACTGAGGAGCTGAGGCAATCGCCGCTTCTCTCTATC 360
Qy      602 TTCCGAGGAGATGATCTGTGGGCACTGGTGAGGCTGTGGCAGTGAAGGCACTGCC 661
Db      361 TTCCGAGGAGATGATCTGTGGGCACTGGTGAGGCTGTGGCAGTGAAGGCACTGCC 420
Qy      662 TGCAGGAGATTTCAACAGCCCAATGCGGAGACCTGCCAGGGGCTCAAGACCAAGAGG 721
Db      421 TGCAGGAGATTTCAACAGCCCAATGCGGAGACCTGCCAGGGGCTCAAGACCAAGAGG 480
Qy      722 ATGTGTTAGCCACGTCAACAACCAAGACATGAGGAGCATGAGCAAGAAAGATCTATGTG 781
Db      481 ATGTGTTAGCCACGTCAACAACCAAGACATGAGGAGCATGAGCAAGAAAGATCTATGTG 540
Qy      782 GCGAGGTTCCTCCGTTAGGCTGTGCTATCCGATATTTGATTTGACGGGATCTCTGGGG 841
Db      541 GCGAGGTTCCTCCGTTAGGCTGTGCTATCCGATATTTGATTTGACGGGATCTCTGGGG 600
Qy      842 TCAACCTGGCGCTGTGGAGCTGTATCTGTGTAAGGCAACAGCGTTGATGCGATAA 901
Db      601 TCAACCTGGCGCTGTGGAGCTGTATCTGTGTAAGGCAACAGCGTTGATGCGATAA 660
Qy      902 CAAGAATCGCTTTAAGTTACAGCAGAGCGCTCAAGAGAGCTCAAGGTGTGCGCGT 961
Db      661 CAAGAATCGCTTTAAGTTACAGCAGAGCGCTCAAGAGAGCTCAAGGTGTGCGCGT 720
Qy      962 GGGCAGCTTTTTCAGACCGGTTCTGTTATACCT-GGGCAGGAGTAGAGACCT-GTGGT 1019
Db      721 GGGCAGCTTTTTCAGACCGGTTCTGTTATACCTGGGAGCAAGGATGATGCTGTGCT 780
Qy      1020 CAAGCATCTTAAGGGCTCTATGAAGAACTAAACAAGAAAGAGGCGGAGCTTCCACCCC 1079
Db      781 CAAGCATCTTAAGGGCTCTATGAAGAACTAAACAAGAAAGGCGGAGGAGGCC 840
```

Qy 1080 AGCCAAACC 1089
Db 841 TGGCAACCC 850

RESULT 10
BI597531
LOCUS
DEFINITION 603247059F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5294607 5',
818 bp mRNA linear EST 07-SEP-2001
BI597531
ACCESSION
VERSION BI597531
KEYWORDS GI:15490470
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 818)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapds-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
CDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM1745 row: f column: 16
High quality sequence stop: 811.
Location/Qualifiers
1..818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5294607"
/issue_type="hypothalamus"
/lab_host="DH10B"
/clone_11b="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamH; Site_2: SalI-XhoI
(gtagag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 37.7%; Score 741.2; DB 3; Length 818;
Blast Local Similarity 95.6%; Pred. No. 3.1e-187;
Matches 783; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

Qy 991 ACCTGGCCAGGAGTGAAGACACTGTGCTCAAGCATCTTAGGGCTCTATGAGAACTA 1050
Db 1 AGCGGGGTGAGGATGAGACACTGTGCTCAAGCATCTTAGGGCTCTATGAGAACTA 60

Qy 1051 ACAAGAAAGCAGGGGAGACTTGCACACCCACCCAAACCCGAGCAAGGGTCTGCGCATCC 1110
Db 61 ACAAGAAAGAGGGGAGACTTGCACACCCACCCAAACCCGAGCAAGGGTCTGCGCATCC 120

Qy 1111 AGGCTGTGCGGCTTCACTGCGGGGAGAAACCTCTGCAAGGGGAGACAGAGCCCTT 1170
Db 121 AGGCTGTGCGGCTTCACTGCGGGGAGAAACCTCTGCAAGGGGAGACAGAGCCCTT 180

Qy 1171 CCAAGCCACACGAGCCGGTCCCGGAGCCATTGTGTCAGCTTCCCGCCACAGGGACACT 1230

Db 181 CAGAGCTTACAGAGCCGGTCCCGGAGCCATTGTGTCAGCTTCCCGCCACAGGGACACT 240

Qy 1231 CGTTCTTCCACACCTGCTCTGCTGGGGGGCTGTCCGGGAAAGACCTACCTCTGGGGACT 1290
Db 241 CGTTCTTCCACACCTGCTCTGCTGGGGGGCTGTCCGGGAAAGACCTACCTCTGGGGACT 300

Qy 1291 CAGGATGTGCGAGCCCGGCTTGGCTCAGGAGAGACCTCAGGATTCCTGAGATTCCTG 1350
Db 301 CAGGATGTGCGAGCCCGGCTTGGCTCAGGAGAGACCTCAGGATTCCTGAGATTCCTG 360

Qy 1351 CAGTGAATCTCATGCCCCCTTCCCAACGAGACTGAGAGTGAAGGCTTGTTCCTG 1410
Db 361 CAGTGAATCTCATGCCCCCTTCCCAACGAGACTGAGAGTGAAGGCTTGTTCCTG 420

Qy 1411 CATTAATATTTGAGAGAGACTGCTGGTCCGTGCCATATCCAGAGAGACAGCTGGCC 1470
Db 421 CATTAATATTTGAGAGAGACTGCTGGTCCGTGCCATATCAAGAGAGACAGCTGGCC 480

Qy 1471 CCTGCTGAGAGGCTGAACACCTGCGGAGCGGGTGAATCGGCTTTCGCTGACCCAGC 1530
Db 481 CCTGCTGAGAGGCTGAACACCTGCGGAGCGGGTGAATCGGCTTTCGCTGACCCAGC 540

Qy 1531 ACTGATTCGAGACAGGGGACACCCCTTCAAGACTAGGAGCAACAGCCGTGTCTCCACC 1590
Db 541 ACTGATTCGAGACAGGGGACACCCCTTCAAGACTAGGAGCAACAGCGTGTCTCCACC 600

Qy 1591 TCAGGGCTTGGCTCTGCGGCGCTCCCACTTGGAAAGTCTCAGTTCCTCCAGGCTTTCAG 1650
Db 601 TCAGGGCTTGGCTCTGCGGCGCTCCCACTTGGAAAGTCTCAGTTCCTCCAGGCTTTCAG 659

Qy 1651 AAGCATCTGGGCGCAGGCGCTCATGAGTGAATTTCCCTAGGCTTAAACAACCCAGCAAG 1710
Db 660 AAGCATCTGGGCGCAGGCGCTCATGAGTGAATTTCCCTAGGCTTAAACAACCCAGCAAG 719

Qy 1711 CTTGCGCTCTCTGTTTA-TTTTGGTTAACTTATGAAATGTATTTAAGAAAGATGCA 1769
Db 720 CTTGCGCATCTCTGTTTATTTTGGTTAACTTATGAAATGTATTTAAGAAAGATGCTG 779

Qy 1770 GCTCGAGAGAGATTCAAGATGGAACACACAGACCCCA 1808
Db 780 AGTCGAGAGAAATTCGAGATGGAACACACAGACCA 818

RESULT 11
BO920917
LOCUS
DEFINITION BO920917 930 bp mRNA linear EST 20-AUG-2002
AGBNCOURT_8926259 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462826
5', mRNA sequence.
BO920917
ACCESSION BO920917.1 GI:22335834
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 930)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM2653 row: f column: 11
High quality sequence stop: 658.
Location/Qualifiers

```

source
1. .930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6462826"
/issue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 101"
/note="Organ: lung; Vector: pOTB1; Site_1: EcorI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcorI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN
Query Match 37.3%; Score 732.6; DB 5; Length 930;
Best Local Similarity 97.2%; Pred. No. 6 6e-185;
Matches 810; Conservative 0; Mismatches 14; Indels 9; Gaps 6;

QY 482 ATGATATCGGAACCAAGACAGCGGAACTACTCCATCTCTTGATATGAGAGATATA 541
DB 1 ATGATATCGGAACCAAGACAGCGGAACTACTCCATCTCTTGATATGAGAGAG-ATA 57
QY 542 ACCCGAGGTGGGCTACTGAGGGAACCTGAGCAATCGCGCTTGTCTCTCTATC 601
DB 58 ACCCGAGGTGGGCTACTGAGGGAACCTGAGCAATCGCGCTTGTCTCTCTATC 117
QY 602 TTCCTGAGGAGTGCATTCTGGGCACTGGTGACGCTGGCAGTGAAGAGCACTCCC 661
DB 118 TTCCTGAGGAGTGCATTCTGGGCACTGGTGACGCTGGCAGTGAAGAGCACTCCC 176
QY 662 TCGAGGATTTTCAACAGCCCAATGCGGGAACCTGCAAGGGGCTTCAAGACCAAGAGC 721
DB 177 TCGAGGATTTTCAACAGCCCAATGCGGGAACCTGCAAGGGGCTTCAAGACCAAGAGC 236
QY 722 ATGTGTATGACAGCTCAACCCCAAGACATGCGGGCATGAGACCAAGAAAGATCTATG 781
DB 237 ATGTGTATGACAGCTCAACCCCAAGACATGCGGGCATGAGACCAAGAAAGATCTATG 296
QY 782 GGCAGTGTCCCGTATGAGCTGCTCATCCGATATTGATGAGCGGATCTCTCGGGC 841
DB 297 GGCAGTGTCCCGTATGAGCTGCTCATCCGATATTGATGAGCGGATCTCTCGGGC 356
QY 842 TCACCTGCGCTGTGAGAGCTGTATCTGTAGAAAGCGCAAGCGCTTGATGATCCGATTA 901
DB 357 TCACCTGCGCTGTGAGAGCTGTATCTGTAGAAAGCGCAAGCGCTTGATGATCCGATTA 416
QY 902 CAAGAATGCGCTTTAAGGTTCAAGAGAGCGCTTCAAGAAAGCTCAAGGTGCGCCGT 961
DB 417 CAAGAATGCGCTTTAAGGTTCAAGAGAGCGCTTCAAGAAAGCTCAAGGTGCGCCGT 476
QY 962 GGGCAGTGTTCGAAACCGGTTCTGTATATCTGGGCGAGGATGAGACATGATGCTCA 1021
DB 477 GGGCAGTGTTCGAAACCGGTTCTGTATATCTGGGCGAGGATGAGACATGATGCTCA 536
QY 1022 AGCATCTTAAAGGCTTATGAAAGAACTAAAGAAAGAGGAGGAGCTTGCCACCCGAG 1081
DB 537 AGCATCTTAAAGGCTTATGAAAGAACTAAAGAAAGAGGAGGAGCTTGCCACCCGAG 596
QY 1082 CCAAAACCGAGCAAGGCTGTGAGCATCAAGCCTGTGCGGCTTCAAGTGGCGGAGAA 1141
DB 597 CCAAAACCGAGCAAGGCTGTGAGCATCAAGCCTGTGCGGCTTCAAGTGGCGGAGAA 656
QY 1142 CCTCTGCAAGGGGAGCAAGGAGCGGCTTCAAGGCGGAGCAAGGCGGCTTCCGCGGCGCA 1201
DB 657 CCTCTGCAAGGGGAGCAAGGAGCGGCTTCAAGGCGGAGCAAGGCGGCTTCCGCGGCGCA 716
QY 1202 TTTGGTCAAGCTTCCCGGAGCGGAGCACTGTCTTCAACACCTCTGC--TGGTGGAGC 1259
DB 717 TTTGGTCAAGCTTCCCGGAGCGGAGCACTGTCTTCAACACCTCTGC--TGGTGGAGC 776

```

```

QY 1260 TGTCCGGGAAGACACTTACC--TGTGGCACTCAGGCTGTGCGGAGCCCGG 1309
DB 777 TGTCCGGGAAGACACTTACCCTGTGGGGAACCTTCAAGGTGTGCGGAGCCCGG 829

RESULT 12
BMS60320 1013 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6564065 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744726
DEFINITION 5', mRNA sequence.
ACCESSION BMS60320
VERSION BMS60320.1 GI:18804652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1013)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Arrayed by: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12766 row: 1 column: 15
High quality sequence start: 36
High quality sequence stop: 745.
Location/Qualifiers
1. 1013
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744726"
/issue_type="medulla"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC library."

ORIGIN
Query Match 37.2%; Score 730.2; DB 3; Length 1013;
Best Local Similarity 92.6%; Pred. No. 3e-184;
Matches 801; Conservative 0; Mismatches 58; Indels 6; Gaps 3;

QY 1 ATGAGAGTGTAGAGTGTGCGGGCAAGTGTGGGCAAGAGCGAGAGACATCATTAAG 60
DB 98 ATGAGAGTGTAGAGTGTGCGGTAGTGTGTGGGCAAGAGCGAGAGACATCATTAAG 157
QY 61 AAATACGAAAGAGGACACCGAGCTGGGCTGCCAGAGCAAGAGGAGCTTAAGCTTTTCA 120
DB 158 AAATACGAAAGAGGACACCGAGCTGGGCTGCCAGAGCAAGAGGAGCTTAAGCTTTTGA 217
QY 121 AGCTACACACACAGTGCATATTGGGATTTGATCATGAGACGAGAGCTGCTCTCTG 180
DB 218 AGCTACACACACAGTGCATATTGGGATTTGATCATGAGACGAGAGCTGCTCTCTG 277
QY 181 ACTGGCGGGAGGCGGAGAAATTGGCGGGGAGATCAGCGAAAGCAAGTGGGTGAT 240

```

Db	278	ACTGGCGCGGAGGTGTAAGCAATTTCGGCGCGGAGTACGCCAAAGACAAAGTGGGTGAAA	337
Qy	241	ATGCTGGGAGACTGGGAGAAATATCAAAAGCAGCAAGAAAGCTCATATGATGAGCGTATCAAG	300
Db	338	ATGCTGGGAGAAATGGGACACTTACAAAACACACAAAGCTCATGATGAGCGTATCAAG	397
Qy	301	GGAAATGCCATATGAATCCGGGGGCCGAAATGATGTGTCATGCTCTCTGAAACATTGAGGAATG	360
Db	388	GGAAATCCCATATGAATCCGGGGGCCGAAATGATGTGTCATGCTCTCTGAAACATTGAGGAATC	457
Qy	361	AAATTGAAAAAACCCCGAAAGATATCAAGATCATGAAAGAAAGGCAAGAGTCATCTGAG	420
Db	458	AAATTGAAAAAACCCCGAAAGATATCAAGATCATGAAAGAAAGGCAAGAGTCATCTGAA	517
Qy	421	CACATCCAGCGCATGACCCGGGACGTAAAGCCGGGACATTAAAGAACATATATCTTTCAG	480
Db	518	CACATCCAGCAATGACCTGAGCGTAAGTGGGACATTAAAGGACATATATCTTTCAGG	577
Qy	481	GATCGATACGGAACCAAGCAGCGGGAACCTATCCATCTCTCTGGCATATGAGGAGTAT	540
Db	578	GATCGATACGGAACCAAGCAGCGGGAACCTATCTCTCTGGCATATGAGGAGTAT	637
Qy	541	AAACCCGAGGTGGCTACTGCAAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT	600
Db	638	AAACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT	697
Qy	601	CTTCCCTGAGAGAGATGATCTTCTGGGACCTGGTGCAGCTGCTGCGCAGTGAAGGACATCC	660
Db	698	CTTCCCTGAGAGAGATGATCTTCTGGGACCTGGTGCAGCTGCTGCGCAGTGAAGGACATCC	757
Qy	661	CTGCAAGGATTTCAAGCCCAATGAGCGGAGCCGTTCAGAGGAGCTTCAAGACCAAGAGAG	720
Db	758	CTGCAAGGATTTCAAGCCCAATGAGCGGAGCCGTTCAGAGGAGCTTCAAGACCAAGAGAG	817
Qy	721	CATGAGTAAAGCCACGTCAACCCAAAGCATGAGGGGCATACAGGACAAAGAAATGTATGT	780
Db	818	CATGAGTAAAGCCACGTCACTACCCAAAGCATACGTGGGCATACAGGACAAAGAAATGTATGT	877
Qy	781	GGGCGAGTGTTC--CCGTTAGGCTGCTCTCATCC--GGAATATGATTGACGGG--ATCTCT	834
Db	878	GGGCGAGTGTGTCTCCCTTAGGCTGCTCATCCGGAATATGATTGAAATGGATCTCTCC	937
Qy	835	CTCGGGCTACACCTCTGCGCTGTGGG	889
Db	938	TCGGGCTACACCTCTGCGCTGTGGG	962
RESULT 13			
B1601698			
LOCUS	B1601698	839 bp	mRNA linear EST 07-SEP-2001
DEFINITION	603249940F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301752 5',		
ACCESSION	B1601698		
VERSION	B1601698.1 GI:15494637		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homidae; Homo.		
	1 (bases 1 to 839)		
REFERENCE	NIH-MGC http://mhc.nci.nih.gov/		
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)		
TITLE	Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Miklos Palciovits, M.D., Ph.D.		
	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizuki		
	Toshiyuki and Piero Carninci (RIKEN)		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		

http://image.llnl.gov
plate: LLML1763 row: p column: 09
High quality sequence scop: 732.
Location/Qualifiers

FEATURES
SOURCE

1. 839
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5301752"
/issue_type="hypochlamus"
/lab_host="DH10B"
/clone_id="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescriptpr (modified
pBluescript KS+); Site 1: BamH; Site 2: SalI-XhoI
(gtagag): Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTTAA-3', size selected for average
insert size 2.3 kb and normalized to R07 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN

Query Match 36.9%; Score 724.8; DB 3; Length 839;
Beet Local Similarity 94.5%; Pred. No. 8e-183;
Matches 794; Conservative 0; Mismatches 42; Indels 4; Gaps 4;

QY 991 ACCTGGCCAGGATGAGGACACGTGCTCAAGATCTTAGGGCTCATGAAACTA 1050
Db 1 AGCGGGGTCAGGACATGAGACATGTGTCAAGATCTTAGGGCTCATGAAACTA 60
QY 1051 ACAAGAAAGCAGGGGGACCTGCGCACCCCCCAACCCGACGAAGGGTCGTGGCATCC 1110
Db 61 ACAAGAAAGAGGGGGACCTGCGCACCCCAACCCGACGAAGGGTCGTGGCATCC 120
QY 1111 AGGCTGTGCGGCTTCACTGTGGGGGAAACCTCTTGCAMGGGGGACAGGACGCCCT 1170
Db 121 AGGCTGTGCGGCTTCACTGTGGGGGAAACCTCTTGCAMGGGGGACAGGACGCCCT 180
QY 1171 CAGAGGCCACAGCCGGGTTCCCGGGGCCATTGTCAGCTTCCCGGACAGGGGACCT 1230
Db 181 CAGAGGCCACAGCCGGGTTCCCGGGGCCATTGTCAGCTTCCCGGACAGGGGACCT 240
QY 1231 CGTCTTCCACACCCTGTCTGTGGGGGCTGTCCGGGAAAGACACTTACCTGTGGGCACT 1290
Db 241 CGTCTTCCACACCCTGTCTGTGGGGGCTGTCCGGGAAAGACACTTACCTGTGGGCACT 300
QY 1291 CAGGGGTGGCCAGCGCCGGCCCTGGGCTCAGGGGAAGACCTCAGGGTTCCTGGAGATCCTG 1350
Db 301 CAGGGGTGGCCAGCGCCGGCCCTGGGCTCAGGGGAAGACCTCAGGGTTCCTGGAGATCCTG 360
QY 1351 CAGTGAAGCTCATGCCCCCGGCTCCCAACGAGACCTGGAAGGGGCTTGGTTCGGC 1410
Db 361 CAGTGAAGCTCATGCCCCCGGCTCCCAACGAGACCTGGAAGGGGCTTGGTTCGGC 420
QY 1411 CATTAATGATTTCAAGACAGAGCTGCTGGGTCCGTGCATATCCACAGAGAGACCACTGGCC 1470
Db 421 CATTAATGATTTCAAGACAGAGCTGCTGGGTCCGTGCATATCCACAGAGAGACCACTGGCC 480
QY 1471 CCGTGTGAGCAGGCTGAACCCCTGCGGAGGGGTGAGATGGGCTTTTCGTGACCCAGC 1530
Db 481 CCGTGTGAGCAGGCTGAACCCCTGCGGAGGGGTGAGATGGGCTTTTCGTGACCCAGC 540
QY 1531 ACTGATTCGACACAGGGGACCCCTTCAGAGCTTAGGGACGAACAGAGTGTGCTCCACCC 1590
Db 541 ACTGATTCGACACAGGGGACCCCTTCAGAGCTTAGGGACGAACAGAGTGTGCTCCACCC 600
QY 1591 TCAGGGGCTTGGCTCTGTGGGCTCCTCACTTGAAGATTCTGATTCCTCCAGGGCTTCAG 1650
Db 601 TCAGGGGCTTGGCTCTGTGGGCTCCTCACTTGAAGATTCTGATTCCTCCAGGGCTTCAG 660
QY 1651 AAGCATCTGGGACAGGGGCTCATGAGCTGAGTAAATTTCCCTAGGCTTAAACCAACGACAG 1710

Db 661 AACGACTGGGCGGAGGCTCATGCTGGATGA-TTCCCTAGGCTTAACAACCAAGAG 719

Qy 1711 CTTCGGCTCTCTCTTTATTTTGGTTAACTATGAATAATGATTAGAA-AGATGCA 1769

Db 720 CTTCGATCTCTCGATTATATGTGGTTAACTTATGAATAAGTCTTTAAGAACAGATGCA 779

Qy 1770 GCTCGA-GAGAGATTGAGA-GATGGAACACACGAGCCCGAGATCACAAGCCACCATG 1827

Db 780 GCTCGAGAGAGATTCCGAGAGTGAACAGACCAAGCCAGATCACAAGACGCTG 839

RESULT 14
BM553146 1141 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6542478 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742850
DEFINITION 5', mRNA sequence.
ACCESSION BM553146
VERSION BM553146.1 GI:18791621
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1141)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL12761 row: k column: 11
High quality sequence stop: 665.
Location/Qualifiers
1. 1141
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5742850"
/issue_type="medulla"
/lab_host="DH10B"
/clone_1b="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 36.6%; Score 718.2; DB 3; Length 1141;
Best Local Similarity 97.5%; Pred. No. 5.1e-181;
Matches 740; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

Qy 1 ATGACGTGTGAGAGTCCGCGGCACTGTGTGGCCACAAGAGGAGAGACATCATTTAG 60

Db 102 ATGACGTGTGAGAGTCCGCGGCACTGTGTGGCCACAAGAGGAGAGACATCATTTAG 161

Qy 61 AATATGAAAAAGGAGACCGAGCTGGGCTGCCAGAGACCAAGGGGCTTAAGCTTTTGA 120

Db 162 AATATGAAAAAGGAGACCGAGCTGGGCTGCCAGAGACCAAGGGGCTTAAGCTTTTGA 221

Qy 121 AGCTACACACACAGCTGATCATTTGGGGATTGTATCATGAGACGAGCTGCTCTG 180

FEATURES

source

Db 222 AGCTACACACACAGCTGATCATTTGGGGATTGTATCATGAGACGAGCTCTCTCTG 281

Qy 181 ACTGCGCGGAGGCGAAGCAATTTGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240

Db 282 ACTGCGCGGAGGCGAAGCAATTTGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 341

Qy 241 ATCTGGGAGACTGGGAGAAATTCAAAGACAGCAAAAGCTCATATGTCAGGCTACAG 300

Db 342 ATCTGGGAGACTGGGAGAAATTCAAAGACAGCAAAAGCTCATATGTCAGGCTACAG 401

Qy 301 GGAATGCCCATGAACATCCGCGGCGCGGATGGTCTCTCTGAAACATTGAGAAATG 360

Db 402 GGAATGCCCATGAACATCCGCGGCGCGGATGGTCTCTCTGAAACATTGAGAAATG 461

Qy 361 AAGTTGAAAAACCCCGAAGATACCAATCATGAGAGAGGCAAGAGTCACTGAG 420

Db 462 AAGATGAAAAACCCCGAAGATACCAATCATGAGAGAGGCAAGAGTCACTGAG 521

Qy 421 CACATCCAGGCTATCGACCGGACCTTAAGCGGACATTAAAGACATATATTTTCAGG 480

Db 522 CACATCCAGGCTATCGACCGGACCTTAAGCGGACATTAAAGACATATATTTTCAGG 581

Qy 481 GATCGATACGAAACCAAGCAGCGGAACTACTACATCTCTCTGGATATGAGAGTAT 540

Db 582 GATCGATACGAAACCAAGCAGCGGAACTACTACATCTCTCTGGATATGAGAGTAT 641

Qy 541 AACCCGAGGTGGCTACTGACAGGACCTGAGCAGATCGCCCTTGTCTCTCTAT 600

Db 642 AACCCGAGGTGGCTACTGACAGGACCTGAGCAGATCGCCCTTGTCTCTCTAT 701

Qy 601 CTTCCTGAGAGATGATCATTTCTGGGCACTGTGTGACCTGTGGCCAGTG-AGAAGCATC 659

Db 702 CTTCCTGAGAGATGATCATTTCTGGGCACTGTGTGACCTGTGGCCAGTGAAGGATTC 761

Qy 660 CCTGCAAGGATTTCAAGGCCCAATAGCCGAGACCGTCCAGGGGCTCAAGACCAAGGA 719

Db 762 CCTGCAAGGATTTCAAGGCCCAATAGCCGAGACCGTCCAGGGGCTCAAGACCAAGGA 821

Qy 720 GCATGTGTAGCCACGTCACAAACCAAGACCATGGGGCA 758

Db 822 GCATGTGTAGCCACCGCTTACCAACCCAGACCTGGGGGA 860

RESULT 15
BI828272 831 bp mRNA linear EST 04-OCT-2001
LOCUS BI828272 603078163F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169999 5',
DEFINITION mRNA sequence.
ACCESSION BI828272
VERSION BI828272.1 GI:15939822
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 831)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1422 row: n column: 16
High quality sequence stop: 822.
Location/Qualifiers

FEATURES

source 1. .831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:516999"
/issue_type="medulla"
/lab_host="DH10B"
/clone_id="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 36.3%; Score 713.6; DB 3; Length 831;
Best Local Similarity 95.4%; Pred. No. 8e-180;
Matches 801; Conservative 0; Mismatches 29; Indels 10; Gaps 6;

```
QY 441 GAGCTAAGCGGGACATTAAAGAAACATATATCTTCAGGGATCGATACGGAAACCAAGA 500
Db 1 GAGCATAGCGGGACATTAAAGAAACATATGTTCTTCAGGGATCGATACGGAAACCAAGA 60
QY 501 GCGGGAACCTACTCCACATCTCCCTGGCATATAGAGAGATTAACCGAGGTGGGCTACTG 560
Db 61 GCGGGAACCTACTCCACATCTCCCTGGCATATAGAGAGATTAACCGAGGTGGGCTACTG 120
QY 561 CAGGACCTAGACACATCGCGGCTTGTCTCTCTATCTTCTGAGAGATGCATT 620
Db 121 CAGGACCTAGACACATCGCGGCTTGTCTCTCTATCTTCTGAGAGATGCATT 180
QY 621 CTGGGCACTGGTGACGTCTGSCCAGTAGAGGCACTCCCTGACAGGATTTCACAGCCC 680
Db 181 CTGGGCACTGGTGACGTCTGSCCAGTAGAGGCACTCCCTGACAGGATTTCACAGCCC 240
QY 681 AAATGGGGGACCCCTCCAGGGGCTCCAAAGCAACAGAGAGATGTGTAGCCAGTCACA 740
Db 241 AAATGGGGGACCCCTCCAGGGGCTCCAAAGCAACAGAGAGATGTGTAGCCAGTCACA 300
QY 741 ACCCAAGACCATGGGGCATCAGACAAGAAAGATCTATGTGGGCAGTTCCTCCGTTAGG 800
Db 301 ACCCAAGACCATGGGGCATCAGACAAGAAAGATCTATGTGGGCAGTTCCTCCGTTAGG 360
QY 801 CTGCTCATCCGATATTTGATTGAAGGAGTCTCTCGGGCTCAACCTGCGCTGTGGGA 860
Db 361 CTGCTCATCCGATATTTGATTGAAGGAGTCTCTCGGGCTCAACCTGCGCTGTGGGA 420
QY 861 CGTGTATCTGTGTAAGGCGCAACAGGCGTGTATGCCGATTAACAAGATCCGCTTAAGGT 920
Db 421 CGTGTATCTGTGTAAGGCGCAACAGGCGTGTATGCCGATTAACAAGATCCGCTTAAGGT 480
QY 921 TCAGCAGAAGCGCTCAGAAAGTCACAGGTGTGSCCGTGGGCAAGTTTTCACACCG 980
Db 481 TCAGCAGAAGCGCTCAGAAAGTCACAGGTGTGSCCGTGGGCAAGTTTTCACACCG 540
QY 981 GTTCGTTGATACCTGGG-CCAGGATGAGACACTGTGCTCAGACATCTTAGGGCTCTTA 1039
Db 541 GTTCGTTGATACCTGGGCGCCAGGATGAGACACTGTGCTCAGACATCTTAGGGCTCTTA 600
QY 1040 TGAAGAAACTTAACAAGAAAGAGAGGGGACCTGCCACCCCGCAACCCGAGCAAGGCT 1099
Db 601 TGAAGAAACTTAACAAGAAAGAGAGGGGACCTGCCA-CCCGAGCAAAACCGAGCAAGGCT 659
QY 1100 CGTGGGATCCAGGCGCTGTGCGGGCTTCACTGTGCGGGAAAGACCTCTGCAAGGGGACA 1159
Db 660 CGTGGGATCCAGGCGCTGTGCGGGCTTCACTGTGCGGGAAAGACCTCTGCAAGGGGAGC 719
QY 1160 GGCAGGCGCCCTCCAGGCGCCACAGCCCGGTTCCGCGGCGCAATTGTGACGTTCCCGC 1219
Db 1160 GGCAGGCGCCCTCCAGGCGCCACAGCCCGGTTCCGCGGCGCAATTGTGACGTTCCCGC 1219
```

```
Db 720 AGGAGGCGCCCTCCAGGCGCCACAGCCCGGTTCCG--GGCAATTGTGACGTTCC--G 774
QY 1220 CAGGGGACCTCGTCTTTCACACCCGTCCTGGTGGGCTGTCCGGGAAGACACTTACC 1279
Db 775 CCAGGGCGCCCTCGTCTTC--CAACCTGTGTGTGGGCTGTCC-GGAGACACTTACC 831
```

Search completed: April 5, 2006, 15:01:52
Job time : 5327 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:18:55 ; Search time 9933 Seconds
(without alignments)
11239.358 Million cell updates/sec

Title: US-10-071-838-1
Perfect score: 1964
Sequence: 1 atggagcgcgtgagagtcgc.....aaatgttcctcgttgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:*

- 1: gb_da:*
- 2: gb_in:*
- 3: gb_env:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pr:*
- 9: gb_ro:*
- 10: gb_str:*
- 11: gb_sy:*
- 12: gb_un:*
- 13: gb_vl:*
- 14: gb_hcg:*
- 15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1964	100.0	1964	6	AX775945	Sequence
2	1964	100.0	1964	8	AF540953	Homo sapi
3	1962.4	99.9	1964	6	AX775947	Sequence
4	1952.8	99.4	2090	8	BC071680	Homo sapi
5	1948	99.2	2072	6	AX086847	Sequence
6	1936.4	98.6	2166	8	BC075809	Homo sapi
7	1902	96.8	2304	6	AX775943	Sequence
8	1900.4	96.8	2304	6	AX775941	Sequence
9	1836.4	93.5	2084	6	AR578516	Sequence
10	1812.4	92.3	2072	6	AR578515	Sequence
11	1809	92.1	2647	6	AX775951	Sequence
12	1807.4	92.0	2647	6	AX775949	Sequence
13	1807.4	92.0	2647	6	AX833552	Sequence
14	1807.4	92.0	2647	8	AK095385	Homo sapi
15	1799.4	91.6	2111	8	BC058890	Homo sapi
16	1753.6	89.3	3389	8	AK122833	Homo sapi
17	1612.4	82.1	1924	8	BC078140	Homo sapi
18	1471.8	74.9	2575	8	BC033670	Homo sapi

19	1220	62.1	7878	6	AX332512 Sequence
20	1220	62.1	7878	8	HSTR210 Homo sapi
21	1218.4	61.8	7971	8	AY143550 Homo sapi
22	1213.6	61.8	8491	8	BSM807865 Homo sapi
23	999	50.9	8201	6	I76205 Sequence 9
24	999	50.9	8201	6	AX774956 Sequence
25	999	50.9	8201	8	HSTR213 Homo sapi
26	882.8	44.9	123925	8	AC003976 Homo sapi
27	879.6	44.8	192881	8	AC067923 Homo sapi
28	878	44.7	37757	8	AC139091 Homo sapi
29	878	44.7	41266	14	AC139176 Homo sapi
30	878	44.7	42073	14	AC144840 Homo sapi
31	878	44.7	89147	14	AC146857 Homo sapi
32	878	44.7	158658	14	AC036181 Homo sapi
33	878	44.7	163628	14	AC139367 Homo sapi
34	878	44.7	185449	14	AC139361 Homo sapi
35	878	44.7	193646	8	AC131056 Homo sapi
36	876.4	44.6	134506	14	AC146482 Homo sapi
37	876.4	44.6	140152	8	AC027821 Homo sapi
38	876.4	44.6	140152	8	AC027821 Homo sapi
39	876.4	44.6	146243	14	AC130293 Homo sapi
40	876.4	44.6	180876	8	AC021317 Homo sapi
41	876.4	44.6	181517	8	AC126327 Homo sapi
42	876.4	44.6	206040	14	AC132814 Homo sapi
43	874.8	44.5	36160	14	AC139089 Homo sapi
44	874.8	44.5	172770	14	AC023133 Homo sapi
45	869.2	44.3	91225	14	AC139556 Homo sapi

ALIGNMENTS

RESULT 1
AX775945
LOCUS AX775945 1964 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 215 from Patent WO03048202.
ACCESSION AX775945
VERSION AX775945.1 GI:32693663

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 Matsuda, A. and Muramatsu, S.
AUTHORS NF-KB activating gene
TITLE Patent: WO 03048202-A 215 12-JUN-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
source location/Qualifiers

CDS
1..1964 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1650 /note="unnamed protein product"
/codon_start=1
/protein_id="CAE11672.1"
/db_xref="GI:32693664"
/translation="MDVVEVAGSWMAQEREDIMKYBKGRAGLPEDKPKPRFSYNN
NVDHGIYHETELPRLTAREAKQIPRAEISRSKQWMDGMEKYSRRKLDRAKGM
PMNIRGPMSTVLINIEKMLKQPRGYQIMKRGKRGSSSHIQRIDVDVGTAKHIFPR
DRYGTQREHLHLILAYEINPEVGYCDLSHIALFLYLPEDBAFVALVOLLSER
HSLOGFHSNGSTVQGLQDQEHVAVTSQPTMGQDKDLCGQSPVAGCLIRLIDG
ISLGLTLRLMNVLYVEGEALMPTIRLAFKVOQKRLTSTSGSPARFCNRFVDWAP
DEDVTLKILRASMKKLATKQGLDLPAPAKEGSSASRVPASRGKTKCKDROAPPG
PPAPRPRIWASPPRARSSTPPCGATVREDITYVGTGQVPSPLACGGPGSGWRFL
QMSGRRLPTLDVGEPRPRHYDRQSCWYVAISGEDDLACQQAHAERVRHAFMA
PSTDSQGTFRARDEQCAPTSGETLGLHLSQFPFG"

ORIGIN

Query Match 100.0%; Score 1964; DB 6; Length 1964;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGAGCTGTATAGAGTCCGGGCAAGTTGGTGGCAACAAGAGAGAGCATCATATG	60	
Db	1	ATGAGCTGTATAGAGTCCGGGCAAGTTGGTGGCAACAAGAGAGCATCATATG	60	
Qy	61	AAATACGAAAAGGAGACACGAGCTGGCTCCAGAGACAAAGGGCTTAAGCTTTTGA	120	
Db	61	AAATACGAAAAGGAGACACGAGCTGGCTCCAGAGACAAAGGGCTTAAGCTTTTGA	120	
Qy	121	AGCTACACAACAACGTCATTTGGGGATTGTATGATGAGAGAGCTCTCTG	180	
Db	121	AGCTACACAACAACGTCATTTGGGGATTGTATGATGAGAGAGCTCTCTG	180	
Qy	181	ACTGCCCGGAGGCGAAGCAATTCGGCGGGAGATCGCCGAAAGACAGTGGTGGAT	240	
Db	181	ACTGCCCGGAGGCGAAGCAATTCGGCGGGAGATCGCCGAAAGACAGTGGTGGAT	240	
Qy	241	ATGCTGGAGACTGGAGAAATACAAAGACAGAAAGCTCATMGATCGAGCTACAG	300	
Db	241	ATGCTGGAGACTGGAGAAATACAAAGACAGAAAGCTCATMGATCGAGCTACAG	300	
Qy	301	GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTCTCTGAACATTTGAGGAATG	360	
Db	301	GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTCTCTGAACATTTGAGGAATG	360	
Qy	361	AAATTGAAAAACCCCGAATAACCATATGAGAGAGAGGCAAGAGTCACTGAG	420	
Db	361	AAATTGAAAAACCCCGAATAACCATATGAGAGAGAGGCAAGAGTCACTGAG	420	
Qy	421	CACATCAGGAGCATCGACCCGGAGCTGAGCGGACATTAAAGAAAGCATATTTTCAG	480	
Db	421	CACATCAGGAGCATCGACCCGGAGCTGAGCGGACATTAAAGAAAGCATATTTTCAG	480	
Qy	481	GATCGATACGAAACCAAGACCGGAGAACTACTCCATCTCTGGCATATGAGAGTAT	540	
Db	481	GATCGATACGAAACCAAGACCGGAGAACTACTCCATCTCTGGCATATGAGAGTAT	540	
Qy	541	AAACCCGAGGTGGCTTACAGAGGACCTGAGCCACATCCCGCTTTTCTCTCTAT	600	
Db	541	AAACCCGAGGTGGCTTACAGAGGACCTGAGCCACATCCCGCTTTTCTCTCTAT	600	
Qy	601	CTTCTGAGAGAGATGATCTGGGCACTGTGCACTGTGGCACTGAGAGCACTCC	660	
Db	601	CTTCTGAGAGAGATGATCTGGGCACTGTGCACTGTGGCACTGAGAGCACTCC	660	
Qy	661	CTGCAAGGATTTTCAAGCCCAATGCGGAGCCGTCAAGGGCTCTCAAGACCAAGAG	720	
Db	661	CTGCAAGGATTTTCAAGCCCAATGCGGAGCCGTCAAGGGCTCTCAAGACCAAGAG	720	
Qy	721	CATGTGTAGCCAGTCAACAACCAAGACATGGGGCATGAGGACAAAGAAATCTATGT	780	
Db	721	CATGTGTAGCCAGTCAACAACCAAGACATGGGGCATGAGGACAAAGAAATCTATGT	780	
Qy	781	GGGCAAGTTCCTCCGTTAGGCTGCTCATCCGATATTTGAATGACGGGATCTCTTCGG	840	
Db	781	GGGCAAGTTCCTCCGTTAGGCTGCTCATCCGATATTTGAATGACGGGATCTCTTCGG	840	
Qy	841	CTCAACCTGCGCTGTGGAGCTGTATCTGTAGAAAGGCAACAGGCTGTGATGCCATA	900	
Db	841	CTCAACCTGCGCTGTGGAGCTGTATCTGTAGAAAGGCAACAGGCTGTGATGCCATA	900	
Qy	901	ACAAGAAATGCCCTTTAAGTTCAAGCAAGCGCTCAAGAAAGTCAAGTGTGGCCG	960	
Db	901	ACAAGAAATGCCCTTTAAGTTCAAGCAAGCGCTCAAGAAAGTCAAGTGTGGCCG	960	
Qy	961	TGGGCACTTTTTCGAACCGGTTCTGTATACCTGGCCAGGAGTGAAGCACTGTGCTC	1020	
Db	961	TGGGCACTTTTTCGAACCGGTTCTGTATACCTGGCCAGGAGTGAAGCACTGTGCTC	1020	
Qy	1021	AAGCATCTTAAAGGCTCTTATGAAAGAACTAAACAAGAAAGAGGGGAGCTTGCCACCCCA	1080	
Db	1021	AAGCATCTTAAAGGCTCTTATGAAAGAACTAAACAAGAAAGAGGGGAGCTTGCCACCCCA	1080	

Qy	1081	GCCAAACCGAAGGAGTCTGCGCATCCAGGCTGTGCCGCTTCACTGCGGGAG	1140	
Db	1081	GCCAAACCGAAGGAGTCTGCGCATCCAGGCTGTGCCGCTTCACTGCGGGAG	1140	
Qy	1141	ACCTCTGCAAGGGGAGCAGAGCGCCCTCCAGGCCCAACAGCCGGTTCCCGGGCC	1200	
Db	1141	ACCTCTGCAAGGGGAGCAGAGCGCCCTCCAGGCCCAACAGCCGGTTCCCGGGCC	1200	
Qy	1201	ATTTGGTCAAGCTTCCCGGCAAGGGACCTGCTTCCACACCTGCTGGTGGGGCT	1260	
Db	1201	ATTTGGTCAAGCTTCCCGGCAAGGGACCTGCTTCCACACCTGCTGGTGGGGCT	1260	
Qy	1261	GTCGGGAAAGACACTACCTGTGGGCACTAGGAGTGTGCCAGCCCGGCTGTGCTAG	1320	
Db	1261	GTCGGGAAAGACACTACCTGTGGGCACTAGGAGTGTGCCAGCCCGGCTGTGCTAG	1320	
Qy	1321	GAGAGACTTCAAGGTTCTTGAGATTCTTGACATGGAATCCATGCCCCGCTCCACAG	1380	
Db	1321	GAGAGACTTCAAGGTTCTTGAGATTCTTGACATGGAATCCATGCCCCGCTCCACAG	1380	
Qy	1381	GACCTGAGATGAGAGGCGCTTGGTCCGCAATTATGATTTGACAGAGCTGTGGTTC	1440	
Db	1381	GACCTGAGATGAGAGGCGCTTGGTCCGCAATTATGATTTGACAGAGCTGTGGTTC	1440	
Qy	1441	CGTGCATATCCAGAGAGACAGCTGAGCCCTGCTGAGAGCTGAACAACCTGCGAG	1500	
Db	1441	CGTGCATATCCAGAGAGACAGCTGAGCCCTGCTGAGAGCTGAACAACCTGCGAG	1500	
Qy	1501	CGGGTGAATCGGCTTTTCCGTGCAACCAGCATGATTTCCGACAGGAGCAACCTTTCAG	1560	
Db	1501	CGGGTGAATCGGCTTTTCCGTGCAACCAGCATGATTTCCGACAGGAGCAACCTTTCAG	1560	
Qy	1561	GCTAAGGAGAAAGAGCGGTGTGCTCCACCTCAGGGGCTTGGCTCGGGGCTCACTTG	1620	
Db	1561	GCTAAGGAGAAAGAGCGGTGTGCTCCACCTCAGGGGCTTGGCTCGGGGCTCACTTG	1620	
Qy	1621	GAAAGTTTCAGTTCCCTCCAGGCTTCTAGAAACATCTGGGCGAGGGCTCATAGCTGAT	1680	
Db	1621	GAAAGTTTCAGTTCCCTCCAGGCTTCTAGAAACATCTGGGCGAGGGCTCATAGCTGAT	1680	
Qy	1681	AAATTCCTTAAAGGCTTAAACCCAGCAAGCTTCCGCTCGTTTATTTTGGTTAA	1740	
Db	1681	AAATTCCTTAAAGGCTTAAACCCAGCAAGCTTCCGCTCGTTTATTTTGGTTAA	1740	
Qy	1741	CTTATGAAATGTATTAAGAAAGTGCAGCTGAGAGAGATTCAAGATGGAACACCC	1800	
Db	1741	CTTATGAAATGTATTAAGAAAGTGCAGCTGAGAGAGATTCAAGATGGAACACCC	1800	
Qy	1801	AGACCCAGATCAACAAGCCACATGAGCCAGCCCTCCAGAGACCCCGCCAGCAGAC	1860	
Db	1801	AGACCCAGATCAACAAGCCACATGAGCCAGCCCTCCAGAGACCCCGCCAGCAGAC	1860	
Qy	1861	CATGTTCTGAATTTGACAGCACCGTGAGCTGCTTGTATCTTAATCTCATGGAAG	1920	
Db	1861	CATGTTCTGAATTTGACAGCACCGTGAGCTGCTTGTATCTTAATCTCATGGAAG	1920	
Qy	1921	ATAACTACCTTCAAGTTTGAATAATGTTTCCCTGTAAGT	1964	
Db	1921	ATAACTACCTTCAAGTTTGAATAATGTTTCCCTGTAAGT	1964	

RESULT 2
AF540953 1964 bp mRNA linear PRI 02-OCT-2002
LOCUS Homo sapiens Rb GTPase-activating protein PRC17 (PRC17) mRNA,
DEFINITION complete cds.
ACCESSION AF540953
VERSION AF540953.1 GI:23452664
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1664)

TITLE

Pei, L., Peng, Y., Ling, X.B., van Eynhoven, W.G., Nguyen, K.C.Q., Rubin, M., Hoey, T., Powers, S. and Li, J.
PRC17, a novel oncogene encoding a Rab GTPase-activating protein, is amplified in prostate cancer

JOURNAL

Cancer Res. (2002) In press

AUTHORS

Li, J.B., Peng, Y. and Powers, S.

REFERENCE

Submitted (23-AUG-2002) Genomics, Tularik, Inc, 266 E Pulaski Road, Suite 1, Greenlawn, NY 11740, USA

TITLE

Location/Qualifiers

JOURNAL

1.1964

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
1.1964
/gene="PRC17"
1.1650
/gene="PRC17"
/codon_start=1
/product="Rab GTPase-activating protein PRC17"
/protein_id="AA033117.1"
/db_xref="GI:2345265"
/translation="MDVAVAGSWMAQEREDIMKYEKGRAGLPEDKPKPRSYNN
NDHGLIVHETELPLTARBAKOIRREISRKSWMDLMDWEKYSRLIIRAYGM
PMNIRGPMWSVLNIEMLKNPGRQIKERKSEHIOIRIDRVSGLTHKHFER
DRYGRKORELHILALAYERNEVGYCRPLSHIALFLIYLPEDAPALYVOLLASER
HSLQGRHSPNGSTVOGLDQOEHVAVATSPKTMGHODKDLGQSGPLCLIRILIDG
ISLGLILRLMDVYLVEGEQALPITRIAKVQOKRLTKSRGPMFRFNKRVDTWAP
DEDTYLKHLRAEMKLTTRKQDLPEPAKEQSSASRPVPAKRGKATLCKGRDMPG
PPARPPRPIWSSPPRAPSSTPCPGAVREDTYPVGTGVSPPALAQGPGQSMRFL
QNSMPRLPTDLDVGGPMFRHYDFRQSCVRAISQEDOLAPCMQAEHPBRVSAFPA
PSTDSQGFPPARBDQPCAPSPGCLGLHSSGFPFGP"

CDS

gene

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

1.1964

1.1650

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 1964; DB 8; Length 1964;

Matches 1964; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGTGTGAGAGTGTGCGGCACTTGTGGGCAAGAGCGAGACATCATTTAG 60
DB 1 ATGACGTGTGAGAGTGTGCGGCACTTGTGGGCAAGAGCGAGACATCATTTAG 60
QY 61 AAATAGGAAAGGAGACCGAGCTGGGCTGCCAGAGGAGCAAGGGGCTTAAGCTTTTGA 120
DB 61 AAATAGGAAAGGAGACCGAGCTGGGCTGCCAGAGGAGCAAGGGGCTTAAGCTTTTGA 120
QY 121 AGCTCAACAAACAGCTGATCATTTGGGGAATTGTATCATAGAGCGAGCTGCTCTCTG 180
DB 121 AGCTCAACAAACAGCTGATCATTTGGGGAATTGTATCATAGAGCGAGCTGCTCTCTG 180
QY 181 ACTGGCGGAGGAGCGGCAAAATTCGGCGGAGATGACCGAAAGCAAGTGGTGAT 240
DB 181 ACTGGCGGAGGAGCGGCAAAATTCGGCGGAGATGACCGAAAGCAAGTGGTGAT 240
QY 241 ATGCTGGGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATATGATGAGCGTATCAAG 300
DB 241 ATGCTGGGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATATGATGAGCGTATCAAG 300
QY 301 GGAATGCGCATGAAATTCGGGCGGCGGATGTGCTCCTCTGAACATTGAGAAATG 360
DB 301 GGAATGCGCATGAAATTCGGGCGGCGGATGTGCTCCTCTGAACATTGAGAAATG 360
QY 361 AAGTTGAAAAACCCGGAAGATACCAATCATGAGAGAGAGGCAAGGTCATCTGAG 420
DB 361 AAGTTGAAAAACCCGGAAGATACCAATCATGAGAGAGAGGCAAGGTCATCTGAG 420
QY 421 CACATTCAGCGCATGACCGGAGCGTAAAGCGGAGCATTAAGGAAGATATATTTCAAG 480

DB 421 CACATTCAGCGCATGACCGGAGCGTAAAGCGGAGCATTAAGGAAGATATATTTCAAG 480
QY 481 GATTCATACGAAACCAAGAGCGGGAATTAATTCATATCTCTGGCATATGAGAGTAT 540
DB 481 GATTCATACGAAACCAAGAGCGGGAATTAATTCATATCTCTGGCATATGAGAGTAT 540
QY 541 AACCGGAGGAGGAGTACGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 600
DB 541 AACCGGAGGAGGAGTACGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 600
QY 601 CTTCCTGAGAGAGATGATCTTGGGCACTGTGTGAGCTCTGTGGCAGTAGAGGACTCC 660
DB 601 CTTCCTGAGAGAGATGATCTTGGGCACTGTGTGAGCTCTGTGGCAGTAGAGGACTCC 660
QY 661 CTGACGAGATTTTCAACGCCAAATGGCGGAGCGTTCAGAGGCTCCAAAGCAAGAG 720
DB 661 CTGACGAGATTTTCAACGCCAAATGGCGGAGCGTTCAGAGGCTCCAAAGCAAGAG 720
QY 721 CATGCTGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 780
DB 721 CATGCTGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 780
QY 781 GGGCAGTGTTCCTGCTTGAAGCTGCTGCTCATCCGGAATTTGATTGACGGAGTCTTCG 840
DB 781 GGGCAGTGTTCCTGCTTGAAGCTGCTGCTCATCCGGAATTTGATTGACGGAGTCTTCG 840
QY 841 CTCACCTGCGCTGCTGAGGAGCGTATCTGTGTAAGAGGAGCAAGGCTTGAATGCCAGTA 900
DB 841 CTCACCTGCGCTGCTGAGGAGCGTATCTGTGTAAGAGGAGCAAGGCTTGAATGCCAGTA 900
QY 901 ACAAAGATGCTCTTTAAGGTTTACAGAGAGCGGCTTCAAGAGCGGAGGAGGAGGAG 960
DB 901 ACAAAGATGCTCTTTAAGGTTTACAGAGAGCGGCTTCAAGAGCGGAGGAGGAGGAG 960
QY 961 TGGGCAAGTGTGTGCAACCGGCTTGTGATACCTGGGCGAGGAGTGAAGACATGCTGCTC 1020
DB 961 TGGGCAAGTGTGTGCAACCGGCTTGTGATACCTGGGCGAGGAGTGAAGACATGCTGCTC 1020
QY 1021 AAGCATTTTGAAGGCTTTATGAAAGAACTTAACAAGAAAGCAAGGGGAGCTTGCACCCCA 1080
DB 1021 AAGCATTTTGAAGGCTTTATGAAAGAACTTAACAAGAAAGCAAGGGGAGCTTGCACCCCA 1080
QY 1081 GCGAAACCGGAGCAAGGAGGCTGCTGAGCTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1140
DB 1081 GCGAAACCGGAGCAAGGAGGCTGCTGAGCTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1140
QY 1141 ACCCTCTGCAAGGAGGAGCAAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1200
DB 1141 ACCCTCTGCAAGGAGGAGCAAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1200
QY 1201 ATTTGGTCAAGCTTCCCGGCAAGGAGCACTGCTTTCACACCTGCTCTGTGGGAGCT 1260
DB 1201 ATTTGGTCAAGCTTCCCGGCAAGGAGCACTGCTTTCACACCTGCTCTGTGGGAGCT 1260
QY 1261 GTTCGGGAGAGCACTTACCTGTGGGCACTCAAGGCTTTCACACCTGCTCTGTGGGAG 1320
DB 1261 GTTCGGGAGAGCACTTACCTGTGGGCACTCAAGGCTTTCACACCTGCTCTGTGGGAG 1320
QY 1261 GTTCGGGAGAGCACTTACCTGTGGGCACTCAAGGCTTTCACACCTGCTCTGTGGGAG 1320
DB 1261 GTTCGGGAGAGCACTTACCTGTGGGCACTCAAGGCTTTCACACCTGCTCTGTGGGAG 1320
QY 1321 GGAAGACCTCAGGAGTTCCTGAGATTCCTGAGTGAAGTCAATGAGGAGGAGGAGGAG 1380
DB 1321 GGAAGACCTCAGGAGTTCCTGAGATTCCTGAGTGAAGTCAATGAGGAGGAGGAGGAG 1380
QY 1381 GACCTGAGAGTGAAGGAGGCTTGTGTTCCGCAATTAATGATTGACAGAGCTGTGGGTC 1440
DB 1381 GACCTGAGAGTGAAGGAGGCTTGTGTTCCGCAATTAATGATTGACAGAGCTGTGGGTC 1440
QY 1441 GGTGCGATATCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 GGTGCGATATCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 CGGAGTGAAGTGGGCTTTCGCTGCAACCGAGCACTGATTCGAGCAAGGAGCCCTTTCAGA 1560

Db 1501 CGGGTGAATCGGCTTTCGTCGACCCGACGATTCGACCAAGGACCCCTTCAGA 1560
Qy 1561 GCTAGGGAAGAACGCGGTGCTCCACCTCAGGGCTTGCTGCGGCTCCACTTG 1620
Db 1561 GCTAGGGAAGAACGCGGTGCTCCACCTCAGGGCTTGCTGCGGCTCCACTTG 1620
Qy 1621 GAAAGTTCGATTCCTCCAGGCTTCTAGAGCATCTGGGCGCAGGCTCATGGCTGAT 1680
Db 1621 GAAAGTTCGATTCCTCCAGGCTTCTAGAGCATCTGGGCGCAGGCTCATGGCTGAT 1680
Qy 1681 AATTTCCCTAGGCTTACCAACCCGACGATTCGCGCTCCCTTTAATTTTGGTTAA 1740
Db 1681 AATTTCCCTAGGCTTACCAACCCGACGATTCGCGCTCCCTTTAATTTTGGTTAA 1740
Qy 1741 CTTATGAAATGTATTAGAAAGATGACGCTCGAGAGATTCAAGATGAAACACACC 1800
Db 1741 CTTATGAAATGTATTAGAAAGATGACGCTCGAGAGATTCAAGATGAAACACACC 1800
Qy 1801 AGACCCGATGACCAAGCCACCATGCGGCTCCGACGACCCCGACCCGACGAC 1860
Db 1801 AGACCCGATGACCAAGCCACCATGCGGCTCCGACGACCCCGACCCGACGAC 1860
Qy 1861 CATGCTTCTGATTCGACGACACCGGAGCTGCTGTTGATCTTAACTCATGAGAG 1920
Db 1861 CATGCTTCTGATTCGACGACACCGGAGCTGCTGTTGATCTTAACTCATGAGAG 1920
Qy 1921 ATAAGTCTTCACTGTTTGAATTAATGTTTCTGTTGAAATG 1964
Db 1921 ATAAGTCTTCACTGTTTGAATTAATGTTTCTGTTGAAATG 1964

RESULT 3
AX775947 1964 bp mRNA linear PAT 14-JUN-2003
LOCUS AX775947
DEFINITION Sequence 217 from Patent WO03048202.
ACCESSION AX775947 GI:32693665
VERSION AX775947.1 GI:32693665
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NP-kB activating gene
JOURNAL Patent: WO 03048202-A 217 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
Source Location/Qualifiers

CDS
1..1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1650
/note="unnamed protein product"
/codon_start=1
/protein_id="CA311673.1"
/db_xref="GI:32693666"
/translation="MDVVEYVAGSWMAQEREDITIMKYEKGRAGLPENKGPPTSTYN
NVHLCIVHETELPLTAREAKQIRREISKRMVDLGMWETKSKRLIDRYKGM
PNIIRGPMWSVLNIEBMKLNKPERVOIMEKGRSESHQRIDRDVSGTLRHRIFR
DRYGTQRELIHLILAYEENPEVGYCRDLSHIALFLVLPREDFAWALVQLASER
HSLOGFSPNGGTVOGLQDOQEHVATSOPTGKHODKDLGCGCSPLGLILIDG
ISLGTLRLMDVYLVEGEOLAMPITTRAFVQOKRLTKTSRCGPMAFCNPFVDTNAR
DEDTVLKHLASMKULTRKQGDLPAPKAPQGSASRPVVASGKTLCKGDQAPG
PPAPFPPIVASAPRAPRSTPCPGAVVEDTVPVGTQVSPBALAGSGPQSMPL
QMSMPLPTDLVEGFWFRRHYDRQSCWTRAIQSDQQLPCWQAEHPARVASAPAA
PSTDSOGTFRARDEQPCAPTSGLCHLESSQGPPEF"

ORIGIN

Query Match 99.9%; Score 1962.4; DB 6; Length 1964;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1963; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGACGTGTGTAGAGGTTCGGGGCACTTGTGTGGGCAAAAGCGAGAGACATCATTTAG 60
1 ATGACGTGTGTAGAGGTTCGGGGCACTTGTGTGGGCAAAAGCGAGAGACATCATTTAG 60
Db 1 ATGACGTGTGTAGAGGTTCGGGGCACTTGTGTGGGCAAAAGCGAGAGACATCATTTAG 60
Qy 61 AATTAAGAAAAGGAGACCGAGCTGGCTGCCAGAGACAAAGGGCTTAAGCTTTTCCA 120
61 AATTAAGAAAAGGAGACCGAGCTGGCTGCCAGAGACAAAGGGCTTAAGCTTTTCCA 120
Db 61 AATTAAGAAAAGGAGACCGAGCTGGCTGCCAGAGACAAAGGGCTTAAGCTTTTCCA 120
Qy 121 AGCTACAAACAACGCTGCATCTTTGGGGAATTGTACATGACGAGAGTCCCTCTG 180
121 AGCTACAAACAACGCTGCATCTTTGGGGAATTGTACATGACGAGAGTCCCTCTG 180
Db 121 AGCTACAAACAACGCTGCATCTTTGGGGAATTGTACATGACGAGAGTCCCTCTG 180
Qy 181 ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCAGCGAAAGCAAGTGGTGGAT 240
181 ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCAGCGAAAGCAAGTGGTGGAT 240
Db 181 ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCAGCGAAAGCAAGTGGTGGAT 240
Qy 241 ATGCTGGAGAGCTGGGAGAAATTCAAAAGCAGCAAAAGCTCATATATGAGGTTCAG 300
241 ATGCTGGAGAGCTGGGAGAAATTCAAAAGCAGCAAAAGCTCATATATGAGGTTCAG 300
Db 241 ATGCTGGAGAGCTGGGAGAAATTCAAAAGCAGCAAAAGCTCATATATGAGGTTCAG 300
Qy 301 GGAATGCCCATGAAACATCCGGGGCCGATGTGTGATGTCCTCTGAAATTTGAAATG 360
301 GGAATGCCCATGAAACATCCGGGGCCGATGTGTGATGTCCTCTGAAATTTGAAATG 360
Db 301 GGAATGCCCATGAAACATCCGGGGCCGATGTGTGATGTCCTCTGAAATTTGAAATG 360
Qy 361 AAGTTGAAAACCCCGGAGATACAGATCATGAGAGAGAGGCAAGATCATCTAG 420
361 AAGTTGAAAACCCCGGAGATACAGATCATGAGAGAGAGGCAAGATCATCTAG 420
Db 361 AAGTTGAAAACCCCGGAGATACAGATCATGAGAGAGAGGCAAGATCATCTAG 420
Qy 421 CACATCAGAGCATTCGACCGGAGCTTAAGCGGAGCATTAAGAGACATATTTCTTCAG 480
421 CACATCAGAGCATTCGACCGGAGCTTAAGCGGAGCATTAAGAGACATATTTCTTCAG 480
Db 421 CACATCAGAGCATTCGACCGGAGCTTAAGCGGAGCATTAAGAGACATATTTCTTCAG 480
Qy 481 GATCGATACGGAACCAAGCAGCGGAGATCTCAATCTCTCTGCAATGAGAGAT 540
481 GATCGATACGGAACCAAGCAGCGGAGATCTCAATCTCTCTGCAATGAGAGAT 540
Db 481 GATCGATACGGAACCAAGCAGCGGAGATCTCAATCTCTCTGCAATGAGAGAT 540
Qy 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATGCGCTTTGTTCTCTCTAT 600
541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATGCGCTTTGTTCTCTCTAT 600
Db 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATGCGCTTTGTTCTCTCTAT 600
Qy 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTGACACTGTCTGGCCAGTGAAGGCACTCC 660
601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTGACACTGTCTGGCCAGTGAAGGCACTCC 660
Db 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTGACACTGTCTGGCCAGTGAAGGCACTCC 660
Qy 661 CTGACAGGATTTTCAAGGCCCAATGCGGAGACCTGACAGGAGCTTCAAGCAACAGAG 720
661 CTGACAGGATTTTCAAGGCCCAATGCGGAGACCTGACAGGAGCTTCAAGCAACAGAG 720
Db 661 CTGACAGGATTTTCAAGGCCCAATGCGGAGACCTGACAGGAGCTTCAAGCAACAGAG 720
Qy 721 CATGTGTAGCCAGCTCACACCCCAAGACCATGAGGATTCAGAGCAAGAAAGTCTATGT 780
721 CATGTGTAGCCAGCTCACACCCCAAGACCATGAGGATTCAGAGCAAGAAAGTCTATGT 780
Db 721 CATGTGTAGCCAGCTCACACCCCAAGACCATGAGGATTCAGAGCAAGAAAGTCTATGT 780
Qy 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATTCGGATATTTGATTAAGCGGATCTCTCGGG 840
781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATTCGGATATTTGATTAAGCGGATCTCTCGGG 840
Db 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATTCGGATATTTGATTAAGCGGATCTCTCGGG 840
Qy 841 CTGACCTGCGCTGTGGAGCGTGTATCTGTGTAAGAGCGAAAGCGGCTTATGATGCGGAT 900
841 CTGACCTGCGCTGTGGAGCGTGTATCTGTGTAAGAGCGAAAGCGGCTTATGATGCGGAT 900
Db 841 CTGACCTGCGCTGTGGAGCGTGTATCTGTGTAAGAGCGAAAGCGGCTTATGATGCGGAT 900
Qy 901 ACAAGATCGCTTTAAGTTTCAGCAAGAGCGCTTCAGAGACGTTCCAGTGTGGCCG 960
901 ACAAGATCGCTTTAAGTTTCAGCAAGAGCGCTTCAGAGACGTTCCAGTGTGGCCG 960
Db 901 ACAAGATCGCTTTAAGTTTCAGCAAGAGCGCTTCAGAGACGTTCCAGTGTGGCCG 960
Qy 961 TGGGCACTGTTTTCGAAACCGGCTTGTGATCTGTGGCCAGGAGTGAAGACATGTGCTC 1020
961 TGGGCACTGTTTTCGAAACCGGCTTGTGATCTGTGGCCAGGAGTGAAGACATGTGCTC 1020
Db 961 TGGGCACTGTTTTCGAAACCGGCTTGTGATCTGTGGCCAGGAGTGAAGACATGTGCTC 1020
Qy 1021 AAGCATCTTAGGGCTCTATGAGAAACTTAAGAAAGAGGAGGAGCTGCAACCCCA 1080
1021 AAGCATCTTAGGGCTCTATGAGAAACTTAAGAAAGAGGAGGAGCTGCAACCCCA 1080
Db 1021 AAGCATCTTAGGGCTCTATGAGAAACTTAAGAAAGAGGAGGAGCTGCAACCCCA 1080

```

Qy 1081 GCCAAACCCGAGCAAGGGTGTGTCGGGATCAGAGCCCTGTCGCGGCTTCACTGAGCGGGAAG 1140
Db 1081 GCCAAACCCGAGCAAGGGTGTGTCGGGATCAGAGCCCTGTCGCGGCTTCACTGAGCGGGAAG 1140
Qy 1141 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCCAGGCGCCACAGCCCGGTTTCCCGCGGCC 1200
Db 1141 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCCAGGCGCCACAGCCCGGTTTCCCGCGGCC 1200
Qy 1201 ATTTGTCAGCTTCCCGGCAAGGGGCACTCGTTTCCACACCCCTGTCCTGTGGGCT 1260
Db 1201 ATTTGTCAGCTTCCCGGCAAGGGGCACTCGTTTCCACACCCCTGTCCTGTGGGCT 1260
Qy 1261 GTCCGGAAGACACCTTACCTTGTGGGCACTCAGAGGTGTGCGCAGCCCGGCTGTGCTAG 1320
Db 1261 GTCCGGAAGACACCTTACCTTGTGGGCACTCAGAGGTGTGCGCAGCCCGGCTGTGCTAG 1320
Qy 1321 GGAGGACCTCAGGGGTTCCCTGGAGATTCTCGAGTGAATCTCATGCCCCCGCTCCCAAG 1380
Db 1321 GGAGGACCTCAGGGGTTCCCTGGAGATTCTCGAGTGAATCTCATGCCCCCGCTCCCAAG 1380
Qy 1381 GACCTGAGACGTAGAGGGGCTTGTGTTCCGSCATTAATTTTCAACAGAGCTGTGGGTC 1440
Db 1381 GACCTGAGACGTAGAGGGGCTTGTGTTCCGSCATTAATTTTCAACAGAGCTGTGGGTC 1440
Qy 1441 CGTGCATATCCCAAGAGAGACAGAGTGGCCCCCTGTGGGAGGCTGAACACCTTGCGGAG 1500
Db 1441 CGTGCATATCCCAAGAGAGACAGAGTGGCCCCCTGTGGGAGGCTGAACACCTTGCGGAG 1500
Qy 1501 CGGGTGAGATGGGCTTTTGGCTGACACCCAGCACTGATTCGAGCAGAGGCAACCCCTTCA 1560
Db 1501 CGGGTGAGATGGGCTTTTGGCTGACACCCAGCACTGATTCGAGCAGAGGCAACCCCTTCA 1560
Qy 1561 GCTAGGGAGCAAGACGCGGTGCTGCCACTTCAAGGGGCTTGTGCTTGGGCTTCACTTG 1620
Db 1561 GCTAGGGAGCAAGACGCGGTGCTGCCACTTCAAGGGGCTTGTGCTTGGGCTTCACTTG 1620
Qy 1621 GAAAGTTCTCAGTTCCTCCAGAGCTTCTAGAAAGCATCTGGGCGAGGCTCATGAGTGAAT 1680
Db 1621 GAAAGTTCTCAGTTCCTCCAGAGCTTCTAGAAAGCATCTGGGCGAGGCTCATGAGTGAAT 1680
Qy 1681 AATTTCCCTAGGCTTTAACAACCCAGCAAGCTTGGCGTCTGTTTATTTTGGTTTAA 1740
Db 1681 AATTTCCCTAGGCTTTAACAACCCAGCAAGCTTGGCGTCTGTTTATTTTGGTTTAA 1740
Qy 1741 CTTATGAAATATGTTATTAAGAAAGATGAGCTCTGAGAGAGATTCAGAGATGAACACACC 1800
Db 1741 CTTATGAAATATGTTATTAAGAAAGATGAGCTCTGAGAGAGATTCAGAGATGAACACACC 1800
Qy 1801 AGACCCCGATCACAAGCAACCATGCCCCGAGCCCTCCAGACCCCGAGCCCAAGAC 1860
Db 1801 AGACCCCGATCACAAGCAACCATGCCCCGAGCCCTCCAGACCCCGAGCCCAAGAC 1860
Qy 1861 CATGCTTCTGAATCTGACGACACCGTGAAGCTGCTTGTATTAATTAATCATGAAGG 1920
Db 1861 CATGCTTCTGAATCTGACGACACCGTGAAGCTGCTTGTATTAATTAATCATGAAGG 1920
Qy 1921 ATAATCACTTCAAGTTTGAATAATATGTTTCTGTTGAATG 1964
Db 1921 ATAATCACTTCAAGTTTGAATAATATGTTTCTGTTGAATG 1964

```

```

RESULT 4
LOCUS BC071680 2090 bp mRNA linear PRI 08-MAR-2005
DEFINITION Homo sapiens TBCL domain family, member 3, mRNA (cDNA clone
ACCESSION BC071680
VERSION BC071680.1 GI:47939522
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

REFERENCE

1 (bases 1 to 2090)
Strusberg RL, Feingold EA, Grouse LH, Derge JC, Klausner RD,
Collins FS, Wagner L, Shennan CM, Schuler GD, Altshul SF, Zeeberg
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Heish F, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
Schetz TE, Brownstein MJ, Uebli TB, Toshitsuki S, Carninci P,
Prange C, Raha SS, Loquellano NA, Peters GJ, Adamson RD, Mullaly
SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Huijck SW,
Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Heltón E, Kerteman M, Madan A, Rodríguez S, Sanchez A, Whiting M,
Madan A, Young AC, Shewchenko Y, Bouffard GG, Blakesley RW,
Touchman JW, Green ED, Dickson WC, Rodríguez AC, Grimwood J,
Schmutz J, Myers RM, Butlerfield YS, Krzywinski MI, Skalka U,
Smallus DE, Scherch A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

CONSRTM

human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

1247932
2 (bases 1 to 2090)

PUBMED

Director MGC Project.
Direct Submission

AUTHORS

Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

JOURNAL

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)

REMARK

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LILN at: <http://image.lln.gov>
Series: IRAX Plate: 166 Row: 9 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1414984.

SOURCE

1..2090
/organism="Homo sapiens"
/mol_type="mRNA"

gene

/db_xref="taxon:9606"
/clone="MGC:87891 IMAG:5165385"
/issue_type="Brain, adult medulla"
/clone_id="NIH.MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

CDS

1..2090
/gene="TBCLD3"
/note="synonyms: PRCL1, TRE17"
/db_xref="GeneID:84218"
/db_xref="MIM:607741"
101..1750
/gene="TBCLD3"
/codon_start=1
/product="TBCL domain family, member 3"
/protein_id="AAH71680.1"
/db_xref="GI:47939523"
/db_xref="GeneID:84218"
/db_xref="MIM:607741"
/translation="MDVVEVAGSWMAQERDITIMYKKGHAGLPEDKGPKPPSYNN
NVHILGIVHETELPLTARAKKRIISRSKRWMLGMWEYKSKSRKLLIDAYKGM
PMNIRGPMVSLNIEKMLKNERYQIMKEKRSSEHQRIDRDISGLRKMPFR

DRYGTORRELLHLLAEBEYNEBVGCRDLISHIALFLYLEBEDAFNALVOLULASER
HSLGPHSHNGVVOGLDQOQKHVAT9PKTMGODKDKDGGQCSPLACLRLIDG
LSLGLTLMLDYLVGEQALMLPTRLIAKVOOKLTTSRGPNARFENRVDPMAR
DEDVTLKILIRASMKULTRKQGLDPPRAKEQSSASRPYPARSGGTLCKEROAPRG
PPARPRLPTWSDSPRAPSSTPCGSAVEDIYVIGTGVSPALAQGPQSMFEL
QNSMPLRPTDLDVBGPMPRHYDFRQSCWVRA1S0BDQALPCWQHPARVRSAPFA
PSTSDQGPFPARDEQCCAPISGPICLGHLHSSQPPPGF"

ORIGIN

Query Match 99.4%; Score 1952.8; DB 8; Length 2090;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1957; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGAGCGTGTGAGGTCGCGGCGAGTGTGTGGGCACAAGCGGAGGACATCATTTATG 60
DB 101 ATGAGCGTGTGAGGTCGCGGCGAGTGTGTGGGCACAAGCGGAGGACATCATTTATG 160
QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGGACAGAGGGGCTTAAGCCTTTTCA 120
DB 161 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGGACAGAGGGGCTTAAGCCTTTTCA 220
QY 121 AGCTACACACACACAGCTCGATCATTTGGGGATTGTACATGAGCGAGCTGCTCTCTG 180
DB 221 AGCTACACACACACAGCTCGATCATTTGGGGATTGTACATGAGCGAGCTGCTCTCTG 280
QY 181 ACTGCGCGGAGGCGAAGCAATTCGCGGGAGATCATGCCGAAAGAGCAAGTGGTGGAT 240
DB 281 ACTGCGCGGAGGCGAAGCAATTCGCGGGAGATCATGCCGAAAGAGCAAGTGGTGGAT 340
QY 241 ATGCTGGGAGACTGGGAAATTAACAAGCAGCAAGAAAGCTCATAGATCGAGCTACAG 300
DB 341 ATGCTGGGAGACTGGGAAATTAACAAGCAGCAAGAAAGCTCATAGATCGAGCTACAG 400
QY 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTGCTCTCTGAAATTTGAGAAATG 360
DB 401 GGAATGCCATGAACATCCGGGGCCCGATGTGTGCTCTCTGAAATTTGAGAAATG 460
QY 361 AAGTTGAAAACCCCGGAGATACAGATCATGAGAGAGAGGAGGAGTGTGATGAG 420
DB 461 AAGTTGAAAACCCCGGAGATACAGATCATGAGAGAGAGGAGGAGTGTGATGAG 520
QY 421 CATCTCAGCGCATTCGACCGGAGCGTAAGCGGAGCATTAAGAAGCATATATTTCTTCA 480
DB 521 CATCTCAGCGCATTCGACCGGAGCGTAAGCGGAGCATTAAGAAGCATATATTTCTTCA 580
QY 481 GATCGATTAAGGAACCAAGCAGCGGAACTACTTCAATCTCTTGGCATATGAGAGTAT 540
DB 581 GATCGATTAAGGAACCAAGCAGCGGAACTACTTCAATCTCTTGGCATATGAGAGTAT 640
QY 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCATCGCGCTTGTCTCTCTAT 600
DB 641 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCATCGCGCTTGTCTCTCTAT 700
QY 601 CTTCCTGAGAGAGTATCTTGTGGGCACTGTGCACTGTGCTGGCCAGTGAAGGCACTCC 660
DB 701 CTTCCTGAGAGAGTATCTTGTGGGCACTGTGCACTGTGCTGGCCAGTGAAGGCACTCC 760
QY 661 CTGCAAGGATTTACAGCCCAAATGGCGGAGCGTCCAGGGGCTCCAAAGCAACAGAG 720
DB 761 CTGCAAGGATTTACAGCCCAAATGGCGGAGCGTCCAGGGGCTCCAAAGCAACAGAG 820
QY 721 CATGTGTGCGCAGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATG 780
DB 821 CATGTGTGCGCAGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATG 880
QY 781 GGGCAGATGTTCCCGTTAGGCTCTCATCCGATATTTGATTGACGGGATCTCTCTCGG 840
DB 881 GGGCAGATGTTCCCGTTAGGCTCTCATCCGATATTTGATTGACGGGATCTCTCTCGG 940
QY 841 CTACACCTGCGCTGTGGAGCTGTATCTGTAGAAAGGCAACAGGCTTGAATGCCGATA 900
DB 941 CTACACCTGCGCTGTGGAGCTGTATCTGTAGAAAGGCAACAGGCTTGAATGCCGATA 1000

QY 901 ACAAGATGCGCTTTAAGGTTCAGAGAAAGCGCTTCACGAGACGTCACAGTGTGGCCG 960
DB 1001 ACAAGATGCGCTTTAAGGTTCAGAGAAAGCGCTTCACGAGACGTCACAGTGTGGCCG 1060
QY 961 TGGGCAAGTTTTTGCACACGGTTCTGTATTAAGTGGCCAGAGGATAGAGACATGCTG 1020
DB 1061 TGGGCAAGTTTTTGCACACGGTTCTGTATTAAGTGGCCAGAGGATAGAGACATGCTG 1120
QY 1021 AAGCATCTTAAGGCGCTCTATGAAGAAATTAACAAGAAAGCAGGGGACCTGCAACCCCA 1080
DB 1121 AAGCATCTTAAGGCGCTCTATGAAGAAATTAACAAGAAAGCAGGGGACCTGCAACCCCA 1180
QY 1081 GCCAAACCGGACCAAGAGTGTGTGATCAGAGCTGTGCTGCGCTTACATGTGCGGAAG 1140
DB 1181 GCCAAACCGGACCAAGAGTGTGTGATCAGAGCTGTGCTGCGCTTACATGTGCGGAAG 1240
QY 1141 ACCCTGTGCAAGGGGACAGGAGGCTCTCAAGGCCCAACAGCCCGTTCCCGCGGCC 1200
DB 1241 ACCCTGTGCAAGGGGACAGGAGGCTCTCAAGGCCCAACAGCCCGTTCCCGCGGCC 1300
QY 1201 ATTTGGTCAAGTTCCTCCGCGCAGGGAAGCTGCTTCCACAAGCTGTCTGTGGGGCT 1260
DB 1301 ATTTGGTCAAGTTCCTCCGCGCAGGGAAGCTGCTTCCACAAGCTGTCTGTGGGGCT 1360
QY 1261 GTCCGGGAAAGACACTTACCTGTGGGACATCAGAGGTGTCCAGCCCGGCTGCTCAG 1320
DB 1361 GTCCGGGAAAGACACTTACCTGTGGGACATCAGAGGTGTCCAGCCCGGCTGCTCAG 1420
QY 1321 GGAAGACCTCAGAGTTCCTGAGATTCCTGCAAGTGAATTCCATGCCCCGCTCCCAAG 1380
DB 1421 GGAAGACCTCAGAGTTCCTGAGATTCCTGCAAGTGAATTCCATGCCCCGCTCCCAAG 1480
QY 1381 GACCTGAGAGTGAAGGAGGCTGTTCCGCAATTAATTAATTTGACAGAGCTGTGGGCT 1440
DB 1481 GACCTGAGAGTGAAGGAGGCTGTTCCGCAATTAATTAATTTGACAGAGCTGTGGGCT 1540
QY 1441 CGTGCCATATCCCAAGAGGACAGAGTGTGCTGTGGCAGGCTGAACAACCTGCGGAG 1500
DB 1541 CGTGCCATATCCCAAGAGGACAGAGTGTGCTGTGGCAGGCTGAACAACCTGCGGAG 1600
QY 1501 CGGATGAGATCGGCTTTCGCTGCAACCAAGCATGANTCCGACAGGGACCCCTTCA 1560
DB 1601 CGGATGAGATCGGCTTTCGCTGCAACCAAGCATGANTCCGACAGGGACCCCTTCA 1660
QY 1561 GCTAGGGAAGAACAGCGCTGTGCTCCACCTCAGGGCTTGTGCTGCGGCTCACTTG 1620
DB 1661 GCTAGGGAAGAACAGCGCTGTGCTCCACCTCAGGGCTTGTGCTGCGGCTCACTTG 1720
QY 1621 GAAAGTTCAGTTCCTCAGGCTTCTAAGAGCATCTGGGCGAGGCTCATGGCTGAT 1680
DB 1721 GAAAGTTCAGTTCCTCAGGCTTCTAAGAGCATCTGGGCGAGGCTCATGGCTGAT 1780
QY 1681 AATTTCCCTAAGCTTAAACAACCAAGCATCTTGGCTCTGTGTTATTTTGGTTAA 1740
DB 1781 AATTTCCCTAAGCTTAAACAACCAAGCATCTTGGCTCTGTGTTATTTTGGTTAA 1840
QY 1741 CTTATGAATGATTTAAGAAAGTGAGCTCGAGAGAGATTTCAAGATGAGAACACACC 1800
DB 1841 CTTATGAATGATTTAAGAAAGTGAGCTCGAGAGAGATTTCAAGATGAGAACACACC 1900
QY 1801 AGACCCAGATCAACAAGCAACATGCCAGGCCCTCCAGACCCCAAGCCCAAGC 1860
DB 1901 AGACCCAGATCAACAAGCAACATGCCAGGCCCTCCAGACCCCAAGCCCAAGC 1960
QY 1861 CATGTTCTGAATTTGAGAGACAGGTGAGCTGTGTTGTATTTAACTATGGAAG 1920
DB 1961 CATGTTCTGAATTTGAGAGACAGGTGAGCTGTGTTGTATTTAACTATGGAAG 2020
QY 1921 ATTAACCTCAAGTTTTGAATTAATGTTTCTGTGAATG 1964
DB 2021 ATTAACCTCAAGTTTTGAATTAATGTTTCTGTGAATG 2064

```
RESULT 5
AX086847
LOCUS AX086847 2072 bp DNA linear PAT 09-MAR-2001
DEFINITION Sequence 799 from Patent WO0112659.
ACCESSION AX086847
VERSION AX086847.1 GI:13276047
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1.
AUTHORS Wiemann, S.
TITLE Human dna sequences
JOURNAL Patent: WO 0112659-A 799 22-FEB-2001;
German Human Genome Project (DE)
FEATURES
source
1..2072
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.2%; Score 1948; DB 6; Length 2072;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 ATGACGTGTGAGAGTCCGCGGCAAGTGGTGGCACAAGAGGAGAGCATCATTAATG 60
DB 99 ATGACGTGTGAGAGTCCGCGGCAAGTGGTGGCACAAGAGGAGAGCATCATTAATG 158
QY 61 AAATACGAAAGGAGACCGAGCTGGGCTGCCAGAGACAGAGGGGCTTAAGCTTTTCCA 120
DB 159 AAATACGAAAGGAGACCGAGCTGGGCTGCCAGAGACAGAGGGGCTTAAGCTTTTCCA 218
QY 121 AGCTACAAACAACGCTCATCTTTGGGGATTGTATCATGAGACGAGAGCTGCTCTG 180
DB 219 AGCTACAAACAACGCTCATCTTTGGGGATTGTATCATGAGACGAGAGCTGCTCTG 278
QY 181 ACTGCGCGGAGCGGAGCAAAATTCGCGCGGAGATCAGCCGAAAGCAAGTGGTGAAT 240
DB 279 ACTGCGCGGAGCGGAGCAAAATTCGCGCGGAGATCAGCCGAAAGCAAGTGGTGAAT 338
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAGCAGACAAAGCTCATTAATGAGGTACAG 300
DB 339 ATGCTGGGAGACTGGGAGAAATACAAAGCAGACAAAGCTCATTAATGAGGTACAG 398
QY 301 GGAATGCCCATGAACATCCGCGGAGCCGATGGGTCACTCTCTGAACAATTGAGAAATG 360
DB 399 GGAATGCCCATGAACATCCGCGGAGCCGATGGGTCACTCTCTGAACAATTGAGAAATG 458
QY 361 AAGTTGAAAAACCCCGGAGATACAGATCATGAAGAGAGGCAAGGATCATCTGAG 420
DB 459 AAGTTGAAAAACCCCGGAGATACAGATCATGAAGAGAGGCAAGGATCATCTGAG 518
QY 421 CACATTCACGCGCATGACCGGAGCGTAAGCGGAGCATTAAGAGAGCATTAATCTTCA 480
DB 519 CACATTCACGCGCATGACCGGAGCGTAAGCGGAGCATTAAGAGAGCATTAATCTTCA 578
QY 481 GATCGATACGGAACCAAGAGCGGAGAACTATCCATCCTCTGCAATATGAGAGTAT 540
DB 579 GATCGATACGGAACCAAGAGCGGAGAACTATCCATCCTCTGCAATATGAGAGTAT 638
QY 541 AACCCGAGAGTGGGCTACTGACGAGGACCTGAGCCATGCGCGCTTGTCTCTCTAT 600
DB 639 AACCCGAGAGTGGGCTACTGACGAGGACCTGAGCCATGCGCGCTTGTCTCTCTAT 698
QY 601 CTTCCTGAGAGAGATGATCTTGGGAGCATGCTGGTCACTGCTGCCAGTGAAGGCACTCC 660
DB 699 CTTCCTGAGAGAGATGATCTTGGGAGCATGCTGGTCACTGCTGCCAGTGAAGGCACTCC 758
QY 661 CTGACGAGGATTCACAGCCCAATGGCGGAGCCGTCAAGGGGCTCCAAGACCAACAGAG 720
```

```
DB 759 CTGACGAGGATTCACAGCCCAATGGCGGAGCCGTCAAGGGGCTCCAAGACCAACAGAG 818
QY 721 CATGTGTAGCCACGCTCAACAACCAAGACATGAGGACATCAGACAAAGAAATCTATGT 780
DB 819 CATGTGTAGCCACGCTCAACAACCAAGACATGAGGACATCAGACAAAGAAATCTATGT 878
QY 781 GGGCAGATGTTCCCGGTTAGGCTGCGCTCATCCGGAATATGACGGGATCTCTCGGG 840
DB 879 GGGCAGATGTTCCCGGTTAGGCTGCGCTCATCCGGAATATGACGGGATCTCTCGGG 938
QY 841 CTCACCCCTGCGCTGTGAGACGTGTATCTGTATGAGAGCGAACAAGGCTGTGATGCCGATA 900
DB 939 CTCACCCCTGCGCTGTGAGACGTGTATCTGTATGAGAGCGAACAAGGCTGTGATGCCGATA 998
QY 901 ACAAGATGCGCTTTAAGGTTACAGAGAACCGCTTACAGAAAGCTGCAAGTGTGCCCCG 960
DB 999 ACAAGATGCGCTTTAAGGTTACAGAGAACCGCTTACAGAAAGCTGCAAGTGTGCCCCG 1058
QY 961 TGGGCACTTTTGGCAACCGGTTCTGTGATCCTGGGCAAGGATGAGACATGTGCTC 1020
DB 1059 TGGGCACTTTTGGCAACCGGTTCTGTGATCCTGGGCAAGGATGAGACATGTGCTC 1118
QY 1021 AAGCATCTTAGGGGCTCTATGAAGAACTAACAAAGAGAGGAGACCTGCAACCCCA 1080
DB 1119 AAGCATCTTAGGGGCTCTATGAAGAACTAACAAAGAGAGGAGACCTGCAACCCCA 1178
QY 1081 GCCAAACCCGAGCAAGGGTGTGTGGCATCAGGCGCTGTGCGGCTTACGTGGCGGAAAG 1140
DB 1179 GCCAAACCCGAGCAAGGGTGTGTGGCATCAGGCGCTGTGCGGCTTACGTGGCGGAAAG 1238
QY 1141 ACCCTTGAAGGGGAGACAGGAGCCCTTCCAGGCTCCACAGCCCGGTTCCCGCGGCC 1200
DB 1239 ACCCTTGAAGGGGAGACAGGAGCCCTTCCAGGCTCCACAGGCGGTTCCCGCGGCC 1298
QY 1201 ATTTGGTCAAGCTTCCCGGCAAGGACCTGCTTTCACACACCTGCTCTGATGGGGCT 1260
DB 1299 ATTTGGTCAAGCTTCCCGGCAAGGACCTGCTTTCACACACCTGCTCTGATGGGGCT 1358
QY 1261 GTCCGGAAAGACACTACCTGTGGCACTCAAGGATGTCAGAGCCCGGCTGAGTCA 1320
DB 1359 GTCCGGAAAGACACTACCTGTGGCACTCAAGGATGTCAGAGCCCGGCTGAGTCA 1418
QY 1321 GAGGACCTCAGGGTCTCTGAGATTCCTGAGATGGAATCTCAATGCCCGGCTCCAAAG 1380
DB 1419 GAGGACCTCAGGGTCTCTGAGATTCCTGAGATGGAATCTCAATGCCCGGCTCCAAAG 1478
QY 1381 GACCTGAGATGAGAGGAGCCCTTGTGCTTCCGCAATTAATGATTTGACAGAGGCTGAGT 1440
DB 1479 GACCTGAGATGAGAGGAGCCCTTGTGCTTCCGCAATTAATGATTTGACAGAGGCTGAGT 1538
QY 1441 CGTGCATATCCACAGAGAGACAGCTGAGCCCTGTGCAAGCTGAACACCTGCGAG 1500
DB 1539 CGTGCATATCCACAGAGAGACAGCTGAGCCCTGTGCAAGCTGAACACCTGCGAG 1598
QY 1501 CGGCTGAGATCGGCTTTGCTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCA 1560
DB 1599 CGGCTGAGATCGGCTTTGCTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCA 1658
QY 1561 GCTAGGAGACGAACAGCGGTGCTCCACCTGAGGAGCCCTGCTGAGGAGCTGCACTTG 1620
DB 1659 GCTAGGAGACGAACAGCGGTGCTCCACCTGAGGAGCCCTGCTGAGGAGCTGCACTTG 1718
QY 1621 GAAAGTTCTCAGTCTCTCAGAGCTTGAAGACATCTGAGGCAAGGCTCATGAGCTGAT 1680
DB 1719 GAAAGTTCTCAGTCTCTCAGAGCTTGAAGACATCTGAGGCAAGGCTCATGAGCTGAT 1778
QY 1681 AATTTCCCTAAGCTTAAACCAAGCAAGCTTGGCTGCTGTTTATTTTGGTTAA 1740
DB 1779 AATTTCCCTAAGCTTAAACCAAGCAAGCTTGGCTGCTGTTTATTTTGGTTAA 1838
QY 1741 CTTATGAAAATGATTAAGAAAGAGTCAAGTCAAGAGATTCAGAGATGAGACACACC 1800
```

Db 1839 CTATGAAATGATTAAGAAAGTGACCTGAGAGATTGAGATGAAACACACC 1898
QY 1801 AGACCCGAGATCAAAAGCCAGCATGCGCCGCTCCGAGCAACCCGACGACAG 1860
Db 1899 AGACCCGAGATCAAAAGCCAGCATGCGCCGCTCCGAGCAACCCGACGACAG 1958
QY 1861 CATGCTTCTGTAATCTGACGACACCGTGAACCTGTTGATTAATCTGATGAAG 1920
Db 1959 CATGCTTCTGTAATCTGACGACACCGTGAACCTGTTGATTAATCTGATGAAG 2018
QY 1921 ATAACTACCTTCACTGATTTGAAATTAATGTTTCTGTTGAATG 1964
Db 2019 ATAACTACCTTCACTGATTTGAAATTAATGTTTCTGTTGAATG 2062

RESULT 6
LOCUS BC075809 2166 bp mRNA linear PRI 17-MAY-2005
DEFINITION Homo sapiens TBC1 domain family, member 3B, mRNA (cDNA clone
MGC:87892 IMAGE:5170064), complete cds.
ACCESSION BC075809
VERSION BC075809.1 GI:49902676
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 2166)
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,K.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uddin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalins,D.E.,
Schurch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 2166)
NIH MGC Project
Direct Submission
Submitted (06-JUL-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>

Series: IRAK Plate: 166 Row: h Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984.
Location/Qualifiers
1. 2166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87892 IMAGE:5170064"
/clone_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
1. 2166
/gene="TBC1D3B"
/note="synonym: PRCL7"
/db_xref="GeneID:414059"
147. 1796
/gene="TBC1D3B"
/codon_start=1
/product="TBC1 domain family, member 3B"
/protein_id="AAH75809.1"
/db_xref="GI:49902677"
/db_xref="GI:49902676"
/db_xref="GeneID:414059"
/translation="MDVAVASWMAQBERDILIMKYEKHRAGLPEDKPKPRSYNN
NVDHLGIHTELEPLTAREAKQIREISRSKVMVLDWERYKSRLIDRAYGM
PMNIRGPMSVILNIEEMKLNKGRQIKMEKGRSEHIQRIIDRISGLRHMFFR
DRYGTQKRELHILAYERNPEVGRDLSHIALFLYLPEBDAFMLVOLASER
HSLQGFSPNSGTVGLODOEHVAVTSQSKTMHGDKDLGCGSLGTLRIILDR
ISGLIRLMDVLYLVEBOLMPTITARKVQOKRTKTSRCRPMARFCNRPDYMAR
DDVTVPKHLNASKKULTRKQGLDPPPAKEQSSASRPVPASRGKTLTGQDQAPG
PPARFPRPSPSPAPASSTPCGAGAVREDYVPGTGAPCPALAOGQDSWPL
QMSMPRLPTLDLVEGPFPHYFROSQWVRALISQDOLAPVQABHPARVARSARA
PGTDSQGPFRARDEQCAPISGFLCGLHLSQFPFPGF"

ORIGIN
Query Match 98.6%; Score 1936.4; DB 8; Length 2166;
Best local Similarity 99.2%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 ATGAGAGTGTATGAGTTCGGGCGGAGTGTGGGCAAGAGGAGGAGCATCATATG 60
Db 147 ATGAGAGTGTATGAGTTCGGGCGGAGTGTGGGCAAGAGGAGGAGCATCATATG 206
QY 61 AAATACGAAAAAGGACACCGAGCTGGGCTCCAGAGACAAAGGCGCTTAAGCTTTTCA 120
Db 207 AAATACGAAAAAGGACACCGAGCTGGGCTCCAGAGACAAAGGCGCTTAAGCTTTTCA 266
QY 121 AGCTACACAAACAGCTGATCATTTGGGAGATTGATCATGAGACGAGAGCTCCTCTG 180
Db 267 AGCTACACAAACAGCTGATCATTTGGGAGATTGATCATGAGACGAGAGCTCCTCTG 326
QY 181 ACTGCGCGGAGGCGGAAAGAAATTGGCGGGAATATGCGGAAAGCAAGTGGTGGAT 240
Db 327 ACTGCGCGGAGGCGGAAAGAAATTGGCGGGAATATGCGGAAAGCAAGTGGTGGAT 386
QY 241 ATGCTGGAGACTGGGAGAAATCAAAAAGCAGAGAAAGCTCATAGTCAGAGCTACAG 300
Db 387 ATGCTGGAGACTGGGAGAAATCAAAAAGCAGAGAAAGCTCATAGTCAGAGCTACAG 446
QY 301 GGAATGCCATGAACATCCGGGCGCGAGTGTGCTCACTCTCTGAACATTAGGAATG 360
Db 447 GGAATGCCATGAACATCCGGGCGCGAGTGTGCTCACTCTCTGAACATTAGGAATG 506
QY 361 AAGTTGAAAAAAGGAGGAGATCCAGATCATGAAGAAGGCAAGAGGTCATCTGAG 420
Db 507 AAGTTGAAAAAAGGAGGAGATCCAGATCATGAAGAAGGCAAGAGGTCATCTGAG 566
QY 421 CACATCAGGCGATCGACCGGAGCTGAAGCGGAGCATTAAGAGGATATATCTTCAG 480
Db 567 CACATCAGGCGATCGACCGGAGCTGAAGCGGAGCATTAAGAGGATATATCTTCAG 626
QY 481 GATGATACGGAACAAAGCAGCGGAACTATCTCAACTCTCTGCGCATATGAGAGTAT 540

Db 627 GATCATACGGAACCAAGACGCGGAACTAATCTCAATCTCTGCGATATGAGAGTAT 686
Qy 541 AACCCGAGAGTGGGCTACTGCAAGGACCTGAGCCACATGCGGCTTGTCTCTCTAT 600
Db 687 AACCCGAGAGTGGGCTACTGCAAGGACCTGAGCCACATGCGGCTTGTCTCTCTAT 746
Qy 601 CTTCGAGAGAGATGATCTTGGGCACTGAGCTGCTGCGGCACTGAGAGGCACTCC 660
Db 747 CTTCGAGAGAGATGATCTTGGGCACTGAGCTGCTGCGGCACTGAGAGGCACTCC 806
Qy 661 CTGACGAGATTTCAAGCCCAATGCGGAGCCGTCGAGGAGCTCCAGACCAACAGAG 720
Db 807 CTGACGAGATTTCAAGCCCAATGCGGAGCCGTCGAGGAGCTCCAGACCAACAGAG 866
Qy 721 CATGAGTATGAGCACTGCAACACCAAGACCATGAGGAGCATGAGCAAGAAATCTATGT 780
Db 867 CATGAGTATGAGCACTGCAACACCAAGACCATGAGGAGCATGAGCAAGAAATCTATGT 926
Qy 781 GGGCAGTGTCCCGGTAGGCTGCGCATCGGATATTGATTGACGGGATCTCTCTCGG 840
Db 927 GGGCAGTGTCCCGGTAGGCTGCGCATCGGATATTGATTGACGGGATCTCTCTCGG 986
Qy 841 CTCAACCTGCGGCTGAGGAGCTGATCTGATGAGAGGCAAGCGCTTGAATGCGGATA 900
Db 987 CTCAACCTGCGGCTGAGGAGCTGATCTGATGAGAGGCAAGCGCTTGAATGCGGATA 1046
Qy 901 ACAAGATATGCTTTAAGTTTCAAGCAAGAGCGCTCAAGAGACGTCCAGGTGCGCG 960
Db 1047 ACAAGATATGCTTTAAGTTTCAAGCAAGAGCGCTCAAGAGACGTCCAGGTGCGCG 1106
Qy 961 TGGGACGTTTTGCAACCGGTTGATGATACCTGAGGAGGAGTATGATCACTGCTC 1020
Db 1107 TGGGACGTTTTGCAACCGGTTGATGATACCTGAGGAGGAGTATGATCACTGCTC 1166
Qy 1021 AAGCATCTTAAAGGCTCTATGAAAGAACTAACAAGAAAGCAAGGAGACCTGCAACCCA 1080
Db 1167 AAGCATCTTAAAGGCTCTATGAAAGAACTAACAAGAAAGCAAGGAGACCTGCAACCCA 1226
Qy 1081 GCCAAACCCGAGCAAGGCTGCTGAGATCAGGAGCTGTCGCGCTTCACTGAGCGGAG 1140
Db 1227 GCCAAACCCGAGCAAGGCTGCTGAGATCAGGAGCTGTCGCGCTTCACTGAGCGGAG 1286
Qy 1141 ACCCTCTGCAAGGAGGAGCAAGGAGCCCTCCAGGAGCCAGCCGCTTCCGCGGCGC 1200
Db 1287 ACCCTCTGCAAGGAGGAGCAAGGAGCCCTCCAGGAGCCAGCCGCTTCCGCGGCGC 1346
Qy 1201 ATTGAGTATGCTTCCCGGCAAGGAGACCTGCTTCTTCAACACCTGCTCTGAGGAGCT 1260
Db 1347 ATTGAGTATGCTTCCCGGCAAGGAGACCTGCTTCTTCAACACCTGCTCTGAGGAGCT 1406
Qy 1261 GTCCGAGAGACACTTACCTGCTGAGCACTAGGAGTGTGCCAGCCCGGCTCTGCTCAG 1320
Db 1407 GTCCGAGAGACACTTACCTGCTGAGCACTAGGAGTGTGCCAGCCCGGCTCTGCTCAG 1466
Qy 1321 GGAAGACCTCAAGGAGTCTCTGAGATTTCTGAGAGTGAATCTCAATGCGGCTCTCCAG 1380
Db 1467 GGAAGACCTCAAGGAGTCTCTGAGATTTCTGAGAGTGAATCTCAATGCGGCTCTCCAG 1526
Qy 1381 GACCTGAGATGAGAGGAGCTTGTGCTTCCGCAATATGATTTTCAAGAGAGTGTGGTTC 1440
Db 1527 GACCTGAGATGAGAGGAGCTTGTGCTTCCGCAATATGATTTTCAAGAGAGTGTGGTTC 1586
Qy 1441 CGTGCATATCTCCAGAGAGCAAGTGTGCTCTGAGAGGCTGAAACCTCTGCGAG 1500
Db 1587 CGTGCATATCTCCAGAGAGCAAGTGTGCTCTGAGAGGCTGAAACCTCTGCGAG 1646
Qy 1501 CGGAGATATGAGGAGTCTTGTGCTGAGCAAGCAATGATTTCCAGAGGAGCAACCTCTTCA 1560
Db 1647 CGGAGATATGAGGAGTCTTGTGCTGAGCAAGCAATGATTTCCAGAGGAGCAACCTCTTCA 1706
Qy 1561 GCTAGGAGAGCAAGAGCGTGTGCTCCACCTGAGGAGCTTGTGCGGCTCTCACTTG 1620

Db 1707 GCTAGGAGAGCAAGAGTGTGCTCCACCTCAAGGAGCTTGTGCGGAGCTCCACTTG 1766
Qy 1621 GAAAGTCTCAAGTCTTCCCTCCAGGCTTTTGAAGCATCTGAGGAGGAGCTATGAGCTGAT 1680
Db 1767 GAAAGTCTCAAGTCTTCCCTCCAGGCTTTTGAAGCATCTGAGGAGGAGCTATGAGCTGAT 1826
Qy 1681 AATTCCCTAGGCTTAAACCAAGCAAGAGCTTGGGCTCTGCTTTTATTTTGGTTAA 1740
Db 1827 AATTCCCTAGGCTTAAACCAAGCAAGAGCTTGGGCTCTGCTTTTATTTTGGTTAA 1886
Qy 1741 CTTATGAAATATGATTAAGAAAGAGTGAAGCTGAGAGAGATTCAGAGATGAGAACACC 1800
Db 1887 CTTATGAAATATGATTAAGAAAGAGTGAAGCTTGAAGACATTCAGAGATGAGAACACC 1946
Qy 1801 AGACCCAGATCAAGAGCCCAAGCATGCGGAGCTTCCAGACACCCCAAGCCCAAGC 1860
Db 1947 AGACCCAGATCAAGAGCCCAAGCATGCGGAGCTTCCAGACACCCCAAGCCCAAGC 2006
Qy 1861 CATGCTGAAATTTGACGACACCGTGAAGCTGCTTGTATCTTAACTATGAGAG 1920
Db 2007 CATGCTGAAATTTGACGACACCTGAGACCTGCTGCTTGTATCTTAACTATGAGAG 2066
Qy 1921 ATAACTACCTTCAAGCTTTTGAATTAATGTTTCTGTTGAA 1962
Db 2067 ATAACTACCTTCAAGCTTTTGAATTAATGTTTCTGTTGAA 2108

RESULT 7
AX775943 2304 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775943
DEFINITION Sequence 213 from Patent WO03048202.
ACCESSION AX775943
VERSION AX775943.1 GI:32693661
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
1 Matsuda, A. and Muramatsu, S.
NF- κ B activating gene
Patent: WO 03048202-A 213 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
1..2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unamed protein product"
/codon_start=1
/protein_id="CAE11671.1"
/db_xref="GI:32693662"

CDS
/translation="MDVVEVSGSWMAOEREDLIIMKEKGRAGLPEDEGPKPRSYNN
NDVHGIYHELEPLTRAREKQIRREISRSKVDIMLDEBKRSKRLIDRAYKM
PMNIRGPMWSVLNLTBEWKLKNPGRYQIMKGRSSSHIRIDRVSGTIRKHLFFR
DRYGTQREHLHLILAEYENPBEVQCRDLSHIALFLYLPBDAFWALVOLASER
HSLQSPSPNGSTVOGLDQDEHVATSPPTMGHODKDLGCGSPGLGILRIDG
ISLGLTILMVLYVEGQALMPITRLAFKVOOKRLTISRGGPARFCNPFIDMAR
DEDVLIHLIRASMKLIRKQGLDLPAPKBEQSSASRPVAPSRGKTLCKDROAPRG
PFGQLPRHGLVLPVPLVSGTPTLMALRVCPARPWLRKEDLRVPEDSCSGTFCPA
SORTWT"

ORIGIN

Query Match 96.8%; Score 1902; DB 6; Length 2304;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1937; Conservative 0; Mismatches 5; Indels 22; Gaps 1;

Qy 1 ATGACGATGATGAGAGTGTGCGGAGCAAGTGTGAGGAGCAAGAGCGAGAGCATATTATG 60
Db 362 ATGACGATGATGAGAGTGTGCGGAGCAAGTGTGAGGAGCAAGAGCGAGAGCATATTATG 421

FEATURES
source Location/Qualifiers
1..2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11670.1"
/db_xref="GI:32693660"
/translation="MDVEASGSWMAQEREDITIMYKEKHALEPEDKPKPRSYNN
NVNHLGIVHETLPLTAREAKQIREISRKSWMDLDEWKYSRLIDRAYKGM
PNVIRGPMWSVLINTEEMKIKNGYQIMKEKGSSEHIDRDSVSTLRKHIFER
DRYGTQKRELHILAYEENPEVGYCRDLSHIALFLYLPEEDFALVOLASER
HSLQGHSPNGGTVQGLQDOQHVVATSDPKTMGHODKDLGGQCSPLGCLIRLLIDG
ISLGLRLRMVDYLVGEBOALMPITRIAKVQOKRLTKTSRGPMPARPCNRVDTYAR
DEDTVLKHLASMKUTRKQGLQPPAKEDGSSASRPVPSRGGTTLCKGROQAPG
PFGQLPRHGLVLPHPVILVSGKPTLMALRVCPARPWLREDLRVPGDSGCTPCPA
SORTWT"

ORIGIN

Query Match 96.8%; Score 1900.4; DB 6; Length 2304;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1936; Conservative 0; Mismatches 6; Indels 22; Gaps 1;

```
QY 1 ATGACGTGTAGAGGTGCGGGCAGTTGTGGGCACAAGAGGAGGACATCATTTG 60
DB 362 ATGACGTGTAGAGGTGCGGGCAGTTGTGGGCACAAGAGGAGGACATCATTTG 421
QY 61 AAATAAGAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCA 120
DB 422 AAATAAGAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCA 481
QY 121 AGCTACAAACAAGCTGATCATTTGGGAGTTGTACATGAGACGAGCTGCTCTCTG 180
DB 482 AGCTACAAACAAGCTGATCATTTGGGAGTTGTACATGAGACGAGCTGCTCTCTG 541
QY 181 ACTGGCGGGGAGCGAAGAAATTCGGGGGAGATCAGCGAAGCAAGTGGGTGAT 240
DB 542 ACTGGCGGGGAGCGAAGAAATTCGGGGGAGATCAGCGAAGCAAGTGGGTGAT 601
QY 241 ATGCTGGAGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATAGATCAGCGTACAG 300
DB 602 ATGCTGGAGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATAGATCAGCGTACAG 661
QY 301 GGAATGCCATGAATCCGGGGCCCGATGTGTCTCTCTGAACTTGAAGAAATG 360
DB 662 GGAATGCCATGAATCCGGGGCCCGATGTGTCTCTCTGAACTTGAAGAAATG 721
QY 361 AAGTTGAAAACCCCGGAGATACAGATCATGAAGGAGAAAGGAGGATCTGTAG 420
DB 722 AAGTTGAAAACCCCGGAGATACAGATCATGAAGGAGAAAGGAGGATCTGTAG 781
QY 421 CACATCCAGCGCATGACCGGAGCGTAAGCGGAGACATTAAGAAAGCATTTATTTCAAG 480
DB 782 CACATCCAGCGCATGACCGGAGCGTAAGCGGAGACATTAAGAAAGCATTTATTTCAAG 841
QY 481 GATCATACGGAACCAAGCAGCGGAACTATCTCAATCTCTGGCATATAGAGAT 540
DB 842 GATCATACGGAACCAAGCAGCGGAACTATCTCAATCTCTGGCATATAGAGAT 901
QY 541 AACCGGAGGTGGGCTATCGAGGAGACCTGAACCATGCGCGCTTGTCTCTCTAT 600
DB 902 AACCGGAGGTGGGCTATCGAGGAGACCTGAACCATGCGCGCTTGTCTCTCTAT 961
QY 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGAGCTGTGCCAGTAGAGGACCTCC 660
DB 962 CTTCTGAGAGAGATGATTTCTGGGCACTGTGAGCTGTGCCAGTAGAGGACCTCC 1021
QY 661 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTTCAGAGGCTCCACAACAGAG 720
DB 1022 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTTCAGAGGCTCCACAACAGAG 1081
```

```
QY 721 CATGTGTAGCAACGTCACAACCAAGACCAATGGGAGCATCAGGACAAAGAAATCTATGT 780
DB 1082 CATGTGTAGCAACGTCACAACCAAGACCAATGGGAGCATCAGGACAAAGAAATCTATGT 1141
QY 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGAGATATTTGACGGGATCTCTCGGG 840
DB 1142 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGAGATATTTGATGACGGGATCTCTCGGG 1201
QY 841 CTCACCTTCGCGCTGGGAGCTGTATCTGTGTGAAGAGGAAACAGGCGTTGATGCGGAT 900
DB 1202 CTCACCTTCGCGCTGGGAGCTGTATCTGTGTGAAGAGGAAACAGGCGTTGATGCGGAT 1261
QY 901 ACAAGATGCGCTTTAAGGTTACAGAGAGCCCTCAACAAGCCTCAGAGTGGCCG 960
DB 1262 ACAAGATGCGCTTTAAGGTTACAGAGAGCCCTCAACAAGCCTCAGAGTGGCCG 1321
QY 961 TGGGACGTTTTTGCACCGGTTCTGTGATCTGGGCAAGGATGAGACATGTGCTC 1020
DB 1322 TGGGACGTTTTTGCACCGGTTCTGTGATCTGGGCAAGGATGAGACATGTGCTC 1381
QY 1021 AAGCATCTTAGGCGCTCTATGAAGAACTTAACAAGAAAGAGGGAGCCTGCACCCCA 1080
DB 1382 AAGCATCTTAGGCGCTCTATGAAGAACTTAACAAGAAAGAGGGAGCCTGCACCCCA 1441
QY 1081 GCCAAACCCGAGCAAGGTCGTGCGGATCAGAGCCCTGTCGCGCTTCACTGGCGGAG 1140
DB 1442 GCCAAACCCGAGCAAGGTCGTGCGGATCAGAGCCCTGTCGCGCTTCACTGGCGGAG 1501
QY 1141 ACCCTCTGCAAGGGGAGCAGGAGGCCCTCCAGGCCCAACAGCCCGGTTCCGCGGCC 1200
DB 1502 ACCCTCTGCAAGGGGAGCAGGAGGCCCTCCAGGCCCAACAGCCCGGTTCCGCGGCC 1539
QY 1201 ATTTGGTCACTTCCCGCCACAGGACCTCGTTCTTCAACCTCTGTCTGGGGCT 1260
DB 1540 ATTTGGTCACTTCCCGCCACAGGACCTCGTTCTTCAACCTCTGTCTGGGGCT 1599
QY 1261 GTCCGGGAAACACCTTACCTGTGGGACTCAGGGGTGGCCAGGCCCGGCTGGCTCAG 1320
DB 1600 GTCCGGGAAACACCTTACCTGTGGGACTCAGGGGTGGCCAGGCCCGGCTGGCTCAG 1659
QY 1321 GGAGAGCTCAGAGGTTCTGTGAGATTCGTGAGTGAATCTCATGCCCGCTCCCAAG 1380
DB 1660 GGAGAGCTCAGAGGTTCTGTGAGATTCGTGAGTGAATCTCATGCCCGCTCCCAAG 1719
QY 1381 GACCTGAGATGAGAGGCCCTTGTGCTGCACTTATGATTTGACAGAGCTGTGGTTC 1440
DB 1720 GACCTGAGATGAGAGGCCCTTGTGCTGCACTTATGATTTGACAGAGCTGTGGTTC 1779
QY 1441 GGTGCATATCCCAAGAGGACAGCTGGCCCTGTGGCAGGCTGAACACCTGGGGAG 1500
DB 1780 GGTGCATATCCCAAGAGGACAGCTGGCCCTGTGGCAGGCTGAACACCTGGGGAG 1839
QY 1501 CGGCTGAGATCGGCTTGTGCTGCAACCAAGCTGATTCGAACAGGACCCCTTCAAG 1560
DB 1840 CGGCTGAGATCGGCTTGTGCTGCAACCAAGCTGATTCGAACAGGACCCCTTCAAG 1899
QY 1561 GCTAAGGACGAACAGCCGTGTCTCCACTCAAGGACCTTGTCTTGGGCTTCACTTG 1620
DB 1900 GCTAAGGACGAACAGCCGTGTCTCCACTCAAGGACCTTGTCTTGGGCTTCACTTG 1959
QY 1621 GAAAGTTCTCAGTTCCCTCCAGGCTTCTGAAGCATCTGGGCAAGGCTCATAGGCTGAT 1680
DB 1960 GAAAGTTCTCAGTTCCCTCCAGGCTTCTGAAGCATCTGGGCAAGGCTCATAGGCTGAT 2019
QY 1681 AATTTCCCTAGGCTTAAACAACCAAGAGCTTCCGCTCTCTGTTTATTTTGGTTAAA 1740
DB 2020 AATTTCCCTAGGCTTAAACAACCAAGAGCTTCCGCTCTCTGTTTATTTTGGTTAAA 2079
QY 1741 CTTATGAAAATGTATTAAGAAAGATGAGCTCGAGAGATTCAGAGATGAGACACCC 1800
DB 2080 CTTATGAAAATGTATTAAGAAAGATGAGCTCGAGAGATTCAGAGATGAGACACCC 2139
QY 1801 AGACCCAGATACAAAGCAACATGCGCAGCCCTCTCCAGACCCCAAGCCCAAGAC 1860
```

|||||
Db 2140 AGACCCAGATCAAAAGCCAAACCATGCCAGCCCTCCAGCAACCCCAAGCCCAAGC 2199
QY 1861 CATCGTTTGAATTCGACGACACCGGAGCTGCGCTTTGACTTTAACTCAATGGAAG 1920
Db 2200 CATCGTTTGAATTCGACGACACCGGAGCTGCGCTTTGACTTTAACTCAATGGAAG 2259
QY 1921 ATAACTACCTTCACGTTTGAATTAATGTTTCTCTGTTGAATG 1964
Db 2260 ATAAACACCTTCATGTTTGAATTAATGTTTCTCTGTTGAATG 2303
RESULT 9
AR578516 2084 bp DNA linear PAT 14-DEC-2004
LOCUS AR578516
DEFINITION Sequence 692 from patent US 6783969.
ACCESSION AR578516
VERSION AR578516.1 GI:56581312
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2084)
AUTHORS Tang, Y.T., Goodrich, R.W., Asundi, V. and Dimaano, R.T.
TITLE Cathepsin V-like polypeptides
JOURNAL Patent: US 6783969-A 692 31-AUG-2004;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
source 1..2084
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 93.5%; Score 1836.4; DB 6; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;
QY 1 ATGAGCGTGTAGAGTTCGCGGCACTTGTGCGACAAAGCGAGGACATCATTTATG 60
Db 163 ATGAGCGTGTAGAGTTCGCGGCACTTGTGCGACAAAGCGAGGACATCATTTATG 222
QY 61 AAATACGAAAGGAGACCGAGCTGGCTCCAGAGGACAAAGGGCTTAAGCCTTTTGA 120
Db 223 AAATACGAAAGGAGACCGAGCTGGCTCCAGAGGACAAAGGGCTTAAGCCTTTTGA 282
QY 121 AGCTACAAACAACGTCGATCATTTGGGATTTGTAATGAGACGAGGCTGCTCTCTG 180
Db 283 AGCTACAAACAACGTCGATCATTTGGGATTTGTAATGAGACGAGGCTGCTCTCTG 342
QY 181 ACTGCGGAGGAGCGAAGCAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240
Db 343 ACTGCGGAGGAGCGAAGCAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 402
QY 241 ATGCTGGAGACTGGGGAATAAAGAGACGAAAGCTCATGATCGAGCTACAG 300
Db 403 ATGCTGGAGACTGGGGAATAAAGAGACGAAAGCTCATGATCGAGCTACAG 462
QY 301 GGAATGCCATGAAATCCCGGGCCCGATGTGTCACTCTCTGAACTATGAGGAATG 360
Db 463 GGAATGCCATGAAATCCCGGGCCCGATGTGTCACTCTCTGAACTATGAGGAATG 522
QY 361 AAGTTGAAAAACCCGGAGATACAGATCATGAAAGAGAGGCAAGGATCATCTGAG 420
Db 523 AAGTTGAAAAACCCGGAGATACAGATCATGAAAGAGAGGCAAGGATCATCTGAG 529
QY 421 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTTAAGAAAGCATATATCTTCAGG 480
Db 530 -ACATCCAGCGCATCGACCGGAGCGTAAGCGGAGCATTTAAGAAAGCATATATCTTCAGG 588
QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCAATCTCTCTGCGATATGAGGAATAT 540
Db 589 GATCGATACGGAACCAAGCAGCGGGAATCTACTCAATCTCTCTGCGATATGAGGAATAT 648

QY 541 AACCCGAGAGTGGGCTACTGACAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT 600
Db 649 AACCCGAGAGTGGGCTACTGACAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT 708
QY 601 CTTTCTGAGAGAGATCTATTCGCGGACATGTGACGCTGTGGCAGTGAAGGCACTCC 660
Db 709 CTTTCTGAGAGAGATCTATTCGCGGACATGTGACGCTGTGGCAGTGAAGGCACTCC 768
QY 661 CTGCAAGGATTTTCAACAGCCCAATGCGGAGACCTGTCAGAGGCTCCAGAACCAAGAG 720
Db 769 CTGCAAGGATTTTCAACAGCCCAATGCGGAGACCTGTCAGAGGCTCCAGAACCAAGAG 828
QY 721 CATGTGTAGCCACGTCACACCCCAAGCATGCGGACATCAGACCAAGAAAGTCTATGT 780
Db 829 CATGTGTAGCCACGTCACACCCCAAGCATGCGGACATCAGACCAAGAAAGTCTATGT 888
QY 781 GGGCAATGTTCCCGGTAGGCTGCTCATCCGGATATGATTTGAGCGGAGTCTCTCGGG 840
Db 889 GGGCAATGTTCCCGGTAGGCTGCTCATCCGGATATGATTTGAGCGGAGTCTCTCGGG 948
QY 841 CTCACCTGCGCTGTGAGACGTGTATCTGTAGAAAGCGAACAAGCGTTGATGCCGAT 900
Db 949 CTCACCTGCGCTGTGAGACGTGTATCTGTAGAAAGCGAACAAGCGTTGATGCCGAT 1008
QY 901 ACAAGATGCGCTTTAAGTTACGACGAAGCCCTTACGAGACGTCAGGTGCGCCG 960
Db 1009 ACAAGATGCGCTTTAAGTTACGACGAAGCCCTTACGAGACGTCAGGTGCGCCG 1068
QY 961 TGGGCAAGTTTGTGCAACCGGTTGTTGATACCTGCGGCAAGGATAGAGCACTGTGCTC 1020
Db 1069 TGGGCAAGTTTGTGCAACCGGTTGTTGATACCTGCGGCAAGGATAGAGCACTGTGCTC 1128
QY 1021 AAGCATCTTAAGGCTCTATGAGAACTTACAAAGAAAGCGAGGACCTGCCACCCCA 1080
Db 1129 AAGCATCTTAAGGCTCTATGAGAACTTACAAAGAAAGCGAGGACCTGCCACCCCA 1188
QY 1081 GCCAAACCGAGACAGGCTGTGTCATCAGGCTGTGCTCGGCTTCACTGTGCGGGAAG 1140
Db 1189 GCCAAACCGAGACAGGCTGTGTCATCAGGCTGTGCTCGGCTTCACTGTGCGGGAAG 1248
QY 1141 ACCCTTGAAGAGGAGGACAGGAGGCGCTCCAGAGCCCAAGCGGTTCCCGGCGCC 1200
Db 1249 ACCCTTGAAGAGGAGGACAGGAGGCGCTCCAGAGCCCAAGCGGTTCCCGGCGCC 1308
QY 1201 ATTTGTCAGCTTCCCGGCAAGGACCTGTTCTTCCACAACCTGTGCTGTGAGGCT 1260
Db 1309 ATTTGTCAGCTTCCCGGCAAGGACCTGTTCTTCCACAACCTGTGCTGTGAGGCT 1368
QY 1261 GTCGAGGAAGACCTTGTGAGCACTCAGGCTGTGCGACAGCCCGGCTGTGCTCAG 1320
Db 1369 GTCGAGGAAGACCTTGTGAGCACTCAGGCTGTGCGACAGCCCGGCTGTGCTCAG 1428
QY 1321 GAGAGACCTCAGGCTTCTGAGAAATCTGCAAGTGAATCTTCAATGACCGGCTCCCAAG 1380
Db 1429 GAGAGACCTCAGGCTTCTGAGAAATCTGCAAGTGAATCTTCAATGACCGGCTCCCAAG 1488
QY 1381 GACCTGAGAGGAGGCTTGTGTTCCGCAATTATGATTTTCAAGAGAGCTGTGAGTTC 1440
Db 1489 GACCTGAGAGGAGGCTTGTGTTCCGCAATTATGATTTTCAAGAGAGCTGTGAGTTC 1548
QY 1441 CGTGCAATATCCAGAGAGACAGCTGAGCCCTGCTGTGAGGCTTAACAACCTGTGAGAG 1500
Db 1549 CGTGCAATATCCAGAGAGACAGCTGAGCCCTGCTGTGAGGCTTAACAACCTGTGAGAG 1608
QY 1501 CGGCTGAGATCGGCTTTCGTGACCCAGACATGATTTCCGACAGAGGACCCGCTTCAG 1560
Db 1609 CGGCTGAGATCGGCTTTCGTGACCCAGACATGATTTCCGACAGAGGACCCGCTTCAG 1668
QY 1561 GCTAAGGAGCAAGCAGCGTGTGCTCCACCTCAGAGGCTTGTGCTGTGCGGCTTCACCTTG 1620
Db 1669 GCTAAGGAGCAAGCAGCGTGTGCTCCACCTCAGAGGCTTGTGCTGTGCGGCTTCACCTTG 1728
QY 1621 GAAAGTTCTCATGTTCCCTCAGGCTTCTAGAAAGCATGTGAGGCAAGGCTCATGAGCTGAGT 1680

Db 1729 GAAAGTTCAGTTCCTCAGGCTTCTGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1788
Qy 1681 AATTTCCCTAAGCTTAAACAACCAAGAGCTTGGCTCTCGTTTATTTTGGTTAAA 1740
Db 1789 AATTTCCCTAAGCTTAAACAACCAAGAGCTTGGCTCTCGTTTATTTTGGTTAAA 1848
Qy 1741 CTTATGAAAATGATTAAGAAAGAGTGCAGCTGAGAGATTCAGAGATGGAACACACC 1800
Db 1849 CTTATGAAAATGATTAAGAAAGAGTGCAGCTGAGAGATTCAGAGATGGAACACACC 1908
Qy 1801 AGACCCAGATCAACAAGCCAAACATGCCAGCCCTCCAGCAGACCCCGAGCCCAAGAC 1860
Db 1909 AGACCCAGATCAACAAGCCAAACATGCCAGCCCTCCAGCAGACCCCGAGCCCAAGAC 1968
Qy 1861 CATCGTCTGAATTTCTGAGCAGACCGGTAGCCTGCTTGTACTTTAACTCAGGAAG 1920
Db 1969 CATCGTCTGAATTTCTGAGCAGACCGGTAGCCTGCTTGTACTTTAACTCAGGAAG 2028
Qy 1921 ATTAACCTTCAAGCTTTGAAATAATGTTTCCTGTAATG 1964
Db 2029 ATTAACCTTCAAGCTTTGAAATAATGTTTCCTGTAATG 2072

RESULT 10
AR578515
LOCUS AR578515 2072 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 691 from patent US 6783969.
ACCESSION AR578515
VERSION AR578515.1 GI:56581311
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2072)
AUTHORS Tang,Y.T., Goodrich,R.W., Auendi,V. and Drmanac,R.T.
TITLE Cathepsin V-like polypeptides
JOURNAL Patent: US 6783969-A 691 31-AUG-2004;
Nuvelo, Inc.; Sunnyvale, CA

FEATURES
source
1..2072
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 92.3%; Score 1812.4; DB 6; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

Qy 1 ATGACGTGTAGAGTCCCGGCACTTGTGGGCAACAAGCGAGAGACATCATTTAG 60
Db 163 ATGACGTGTAGAGTCCCGGCACTTGTGTGGGCAACAAGCGAGAGACATCATTTAG 222
Qy 61 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAGGGGCTTAAGCTTTTGA 120
Db 223 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAGGGGCTTAAGCTTTTGA 282
Qy 121 AGCTACAAACAACGTCGATCATTTGGGGAATTGATCATGAGACGAGAGCTGCTCTCG 180
Db 283 AGCTACAAACAACGTCGATCATTTGGGGAATTGATCATGAGACGAGAGCTGCTCTCG 342
Qy 181 ACTGGCGGGGAGCGCAAGCAATTCGGCGGGAGATAGCGCAAAAGCAAGTGGTGGAT 240
Db 343 ACTGGCGGGGAGCGCAAGCAATTCGGCGGGAGATAGCGCAAAAGCAAGTGGTGGAT 402
Qy 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAACTCATAGATCGAGCTGCAAG 300
Db 403 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAACTCATAGATCGAGCTGCAAG 462
Qy 301 GGAATGCCATGAACATCCGGGGCCCGGATGTGCTCAGTCTCTGAACTTGAAGAAATG 360
Db 463 GGAATGCCATGAACATCCGGGGCCCGGATGTGCTCAGTCTCTGAACTTGAAGAAATG 522

Qy 361 AAGTGAAAAACCCGGAAGATACAGATCATGAAGAGAGCAAGAGTCACTGAG 420
Db 523 AAGTGAAAAACCCGGAAGATACAGATCATGAAGAGAGCAAGAGTCACTGAG 582
Qy 421 CAATTCAGCGATGACCGGGAGCGTAAGCGGGACAATTAAAGAAAGATATATCTTCAG 480
Db 583 CAATTCAGCGATGACCGGGAGCGTAAGCGGGACAATTAAAGAAAGATATATCTTCAG 642
Qy 481 GATCGATACGAAACCAAGCAGCGGGAATCTATCTCAGATCTCTGGCATATGAGAGTAT 540
Db 643 GATCGATACGAAACCAAGCAGCGGGAATCTATCTCAGATCTCTGGCATATGAGAGTAT 702
Qy 541 AACCGGAGTGGGCTACTGAGGAGACCTGAGCAATGCGCGCTGTTCCTCTCAT 600
Db 703 AACCGGAGTGGGCTACTGAGGAGACCTGAGCAATGCGCGCTGTTCCTCTCAT 762
Qy 601 CTTCTGAGAGAGATGATCTTGGGCACTGATGAGCTGCTGCGCAGTGAAGGCACTCC 660
Db 763 CTTCTGAGAGAGATGATCTTGGGCACTGATGAGCTGCTGCGCAGTGAAGGCACTCC 822
Qy 661 CTGCAAGGATTTCAAGCCCAATGGCGGAGACCTGTCAGGGGCTCCAAACCAAGAG 720
Db 823 CTGCAAGGATTTCAAGCCCAATGGCGGAGACCTGTCAGGGGCTCCAAACCAAGAG 882
Qy 721 CATGTGTAGCCAGCTCAACCAACCAAGCATGGGGCATCAGGACAAAGAAATCATGT 780
Db 883 CATGTGTAGCCAGCTCAACCAACCAAGCATGGGGCATC----- 923
Qy 781 GGGCAGTGTTCCTCGTAAAGCTGCTCATCCGGAATGATTTGACGGGATCTCTCGGG 840
Db 924 -----GATCTCTCTCGGG 936
Qy 841 CTCACCTGCGCTGCTGGAGCTGTATCTGTGAAGAGCGCAACAGGCTTGATGCCGAT 900
Db 937 CTCACCTGCGCTGCTGGAGCTGTATCTGTGAAGAGCGCAACAGGCTTGATGCCGAT 996
Qy 901 ACAAGAATGCGCTTAAGGTTACAGAGAGGCGCTCAGAGAAACGTCAGAGTGGCCGG 960
Db 997 ACAAGAATGCGCTTAAGGTTACAGAGAGGCGCTCAGAGAAACGTCAGAGTGGCCGG 1056
Qy 961 TGGGCACTTTTGGCAACCGGTTCTGTGTAATCTGCGGCGAGGATGAGAGCACTGCTC 1020
Db 1057 TGGGCACTTTTGGCAACCGGTTCTGTGTAATCTGCGGCGAGGATGAGAGCACTGCTC 1116
Qy 1021 AAGCATTTAAGGCTCTATGAAGAACTAACAGAAACAGAGGGGACCTGCCACCCCA 1080
Db 1117 AAGCATTTAAGGCTCTATGAAGAACTAACAGAAACAGAGGGGACCTGCCACCCCA 1176
Qy 1081 GCCAAACCGGACAGAGGTCGTGCGCATCCAGGCGCTGAGCGGGCTTCAAGTGGCGGAAG 1140
Db 1177 GCCAAACCGGACAGAGGTCGTGCGCATCCAGGCGCTGAGCGGGCTTCAAGTGGCGGAAG 1236
Qy 1141 ACCCTCTGCAAGGGGGAAGAGGAGGCGCTCCAGGCGCAACAGCCCGGTTCCCGGGCCC 1200
Db 1237 ACCCTCTGCAAGGGGGAAGAGGAGGCGCTCCAGGCGCAACAGCCCGGTTCCCGGGCCC 1296
Qy 1201 AATTGCTAAGCTTCCCGGCAAGGCACTGCTTCTTCCACACCTGCTCTGTGGGAGCT 1260
Db 1297 AATTGCTAAGCTTCCCGGCAAGGCACTGCTTCTTCCACACCTGCTCTGTGGGAGCT 1356
Qy 1261 GTCCGGGAAGACACTTACCTGTGGGCACTAGGGGTGGGCCAGGCCCGGCTGTGCTAG 1320
Db 1357 GTCCGGGAAGACACTTACCTGTGGGCACTAGGGGTGGGCCAGGCCCGGCTGTGCTAG 1416
Qy 1321 GGAGAGCTCAGGGGTTCTGTGAGATTCGTGAGTGAATCCAGTCCCGCTCCCAAG 1380
Db 1417 GGAGAGCTCAGGGGTTCTGTGAGATTCGTGAGTGAATCCAGTCCCGCTCCCAAG 1476
Qy 1381 GACTTGAACGTAGAGGGCTTGTGTTCCGCCATTATGATTTGACAGAGCTGCTGGGTC 1440
Db 1477 GACTTGAACGTAGAGGGCTTGTGTTCCGCCATTATGATTTGACAGAGCTGCTGGGTC 1536
Qy 1441 CGTGCATATCCCAAGAGAGACAGCTGGCCCCCTGCTGAGGCTGAAACACCTTGGGAG 1500

Db 1537 CGTCCATATCCGAGGAGACAGCTGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1596
Qy 1501 CGGGTGGATCGGCTTTCGCTGCAACCAAGCATGATCCGACCAAGGACACCCCTTGA 1560
Db 1597 CGGGTGGATCGGCTTTCGCTGCAACCAAGCATGATCCGACCAAGGACACCCCTTGA 1566
Qy 1561 GCTAGGAGCAACAGCGGTGTCTCCACCTCAGGGCTTTCCTGCGGCTTCACTTG 1620
Db 1657 GCTAGGAGCAACAGCGGTGTCTCCACCTCAGGGCTTTCCTGCGGCTTCACTTG 1716
Qy 1621 GAAAGTTCAGTTCCTCCAGGCTTTCAGAGCATTTGGGCAAGGCTCATGCTGAT 1680
Db 1717 GAAAGTTCAGTTCCTCCAGGCTTTCAGAGCATTTGGGCAAGGCTCATGCTGAT 1776
Qy 1681 AATTTCCCTAGGCTTTCAGAGCATTTGGGCAAGGCTCATGCTGAT 1740
Db 1777 AATTTCCCTAGGCTTTCAGAGCATTTGGGCAAGGCTCATGCTGAT 1836
Qy 1741 CTTATGAAATGTATTAAGAAAGATGCACTGAGAGATTCAGAGATGAACACACC 1800
Db 1837 CTTATGAAATGTATTAAGAAAGATGCACTGAGAGATTCAGAGATGAACACACC 1896
Qy 1801 AGACCCGAGATCACAAGACCAACCATGCCCCCTCCAGCAACCCCAAGCCACGAC 1860
Db 1897 AGACCCGAGATCACAAGACCAACCATGCCCCCTCCAGCAACCCCAAGCCACGAC 1956
Qy 1861 CATCGTTCTGAATTCAGACCAACCGTGAAGCTGCTTGTGATTTAACTCATGAGAG 1920
Db 1957 CATCGTTCTGAATTCAGACCAACCGTGAAGCTGCTTGTGATTTAACTCATGAGAG 2016
Qy 1921 ATAACTACTTCAAGTTTGAATAATGTTTCTGTGAATG 1964
Db 2017 ATAACTACTTCAAGTTTGAATAATGTTTCTGTGAATG 2060

RESULT 11
AX775951 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775951
DEFINITION Sequence 221 from Patent WO03048202.
ACCESSION AX775951
VERSION AX775951.1 GI:32693669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 221 12-JUN-2003;
Abehi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. .2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB11675.1"
/translation="MDYRVVAGSWMAOERDIIIMKYKGRAGLPENKGPSPSYNN
NVDHGLIVHTEPLPLTARAKQIRRIISKSKVMDLGMWERTKSRKIDAYKRM
PMNIRGPMWSVLNTEEMKLNKPERYQIMKEGRSSEHIQRIDRVSGLTRKHIFPR
DRYTKORELHITLALAYBYNPEVGYCRDISHIALFLYLPEDAFMALVOLIASBR
HSIQGFSPNGSTVOGLQDOOEHVATSRDPTMWHOTSLGLTLMLMVLVEGOALM
PITRIAKVOOKLTKTSRCGPMAFPCNRFDVTWABEDTVLKLHLSMKLTKRQSD
LBPAPKBOSSASARPVPASRGKTLCKGRORAPGCPAPFPPIWASPPRAPRSST
PCRGAVREBTYTVGTGVSPALAOGPGSWRFLOWNSMPLPTDLVDEGWFPHY
DFROSQWRAISQEDQLPCWQAEHPAEVRSAFAASTDSQGTTPRADDEPCAPT
SGPCLGHLHSSQFPFGP"

ORIGIN
Query Match 92.1%; Score 1809; DB 6; Length 2647;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 5; Indels 66; Gaps 1;
Qy 1 ATGAGCGTGTAGAGGTCCGGGACGTTGGTGGCAACAGGCGAGAGCATATTATG 60
Db 755 ATGAGCGTGTAGAGGTCCGGGACGTTGGTGGCAACAGGCGAGAGCATATTATG 814
Qy 61 AATAAGAAAAGGACACCGAGCTGGCTCCAGAGACAAAGGGCTTAAGCTTTTGA 120
Db 815 AATAAGAAAAGGACACCGAGCTGGCTCCAGAGACAAAGGGCTTAAGCTTTTGA 874
Qy 121 AGCTACAAACAACAGCTGCATCTTTGGGGAATTGTACATGACGAGGCTGCTCTG 180
Db 875 AGCTACAAACAACAGCTGCATCTTTGGGGAATTGTACATGACGAGGCTGCTCTG 934
Qy 181 ACTGCGCGGAGGCGAAGCAAAATTGCGGAGATCAGCCGAAAGACAAAGTGTGAT 240
Db 935 ACTGCGCGGAGGCGAAGCAAAATTGCGGAGATCAGCCGAAAGACAAAGTGTGAT 994
Qy 241 ATGCTGGAGACTGGAGAAATTCAAAAGCAGCAAAAGCTCATAGTCAAGGTACAG 300
Db 995 ATGCTGGAGACTGGAGAAATTCAAAAGCAGCAAAAGCTCATAGTCAAGGTACAG 1054
Qy 301 GGAATGCCCATGAACATCCGGGAGCCGATGTGTCAAGTCTCTGAACATTTGAGAAATG 360
Db 1055 GGAATGCCCATGAACATCCGGGAGCCGATGTGTCAAGTCTCTGAACATTTGAGAAATG 1114
Qy 361 AAGTTGAAAACCCCGGAAAGATTCAGATCATGAAGAGAGGCAAGAGTCATCTAG 420
Db 1115 AAGTTGAAAACCCCGGAAAGATTCAGATCATGAAGAGAGGCAAGAGTCATCTAG 1174
Qy 421 CACATCGAGGCAATCCAGCGGAGCTTAAGGAGCAATTAAGAGCATATTCTTCAGG 480
Db 1175 CACATCGAGGCAATCCAGCGGAGCTTAAGGAGCAATTAAGAGCATATTCTTCAGG 1234
Qy 481 GATCGATACGGAACCAAGCAGCGGAACTACTCACAATCTCTGCGATATGAGAGTAT 540
Db 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCACAATCTCTGCGATATGAGAGTAT 1294
Qy 541 AACCCGAGGTGGGCTACTGACAGGAGCTTGAGCCACATGCGCCCTTGTCTCTAT 600
Db 1295 AACCCGAGGTGGGCTACTGACAGGAGCTTGAGCCACATGCGCCCTTGTCTCTAT 1354
Qy 601 CTTCTGAGAGAGATGATCTTCTGGGCACTGTGTCACTGTGGCCAGTGAAGCACTCC 660
Db 1355 CTTCTGAGAGAGATGATCTTCTGGGCACTGTGTCACTGTGGCCAGTGAAGCACTCC 1414
Qy 661 CTGCAAGGATTTTCAAGCCCAATGAGCGGAGCCGTCCAGGGGCTCCAAAGCAAGAGAG 720
Db 1415 CTGCAAGGATTTTCAAGCCCAATGAGCGGAGCCGTCCAGGGGCTCCAAAGCAAGAGAG 1474
Qy 721 CATGTGTAGCCAGCTCACAACCCCAAGACCATGAGGAGCATCAGAACAAAGATCTATGT 780
Db 1475 CATGTGTAGCCAGCTCACAACCCCAAGACCATGAGGAGCATC----- 1515
Qy 781 GGGCAGTGTCCCCGTTAGGCTCCTCATCCGGAATTGATTTGACGGGATCTCTCGGG 840
Db 1516 -----GATCTCTCGGG 1528
Qy 841 CTCACCTGCGCTGTGGAGCGTGTACTGTGTAAGAGGAGCAAGGCGTTGATGCGGATA 900
Db 1529 CTCACCTGCGCTGTGGAGCGTGTACTGTGTAAGAGGAGCAAGGCGTTGATGCGGATA 1588
Qy 901 ACAAGATGCGCTTTAAGTTTCAGAGAAAGCGCTTCAGAGAGCGTCCAGGTGCGCG 960
Db 1589 ACAAGATGCGCTTTAAGTTTCAGAGAAAGCGCTTCAGAGAGCGTCCAGGTGCGCG 1648
Qy 961 TGGGCACTTTTTCAGACCGGTTCTGTGATACCTGGGCGCAGGAGTAGAGCACTGTGCTC 1020
Db 1649 TGGGCACTTTTTCAGACCGGTTCTGTGATACCTGGGCGCAGGAGTAGAGCACTGTGCTC 1708

1021 AAGCATCTTAGGGCCCTCTATGAAAGAACTAACAAAGAGAGGGGACCTGCAACCCCA 1080
1709 AAGCATCTTAGGGCCCTCTATGAAAGAACTAACAAAGAGAGGGGACCTGCAACCCCA 1768
1081 GCCAAACCCGAGCAAGGGTCTGTGGGCACTCAAGGCTGTGCGGCTTACGTGGCGGAAG 1140
1769 GCCAAACCCGAGCAAGGGTCTGTGGGCACTCAAGGCTGTGCGGCTTACGTGGCGGAAG 1828
1141 ACCCTCTGCAAGGGGAGACAGGAGGCGCTTCCAGGCGCCACAGCCCGGTTCCGCGGCC 1200
1829 ACCCTCTGCAAGGGGAGACAGGAGGCGCTTCCAGGCGCCACAGCCCGGTTCCGCGGCC 1888
1201 ATTGGTCAAGCTTCCCGGCAAGGGACCTGCTTCCACACCTGCTGGTGGGGCT 1260
1889 ATTGGTCAAGCTTCCCGGCAAGGGACCTGCTTCCACACCTGCTGGTGGGGCT 1948
1261 GTCCGGGAGACACCTACCTGTGGCACTCAAGGAGTGTGCCAGCCCGGCTGGCTCAG 1320
1949 GTCCGGGAGACACCTACCTGTGGCACTCAAGGAGTGTGCCAGCCCGGCTGGCTCAG 2008
1321 GGAGAACCTCAGGGTCTCTGAGATTCTGCAAGTGAATCCATGCCCGGCTCCCAAG 1380
2009 GGAGAACCTCAGGGTCTCTGAGATTCTGCAAGTGAATCCATGCCCGGCTCCCAAG 2068
1381 GACCTGAGCTAGAGGGGCTTGGTTCGCGCAATTAATGATTTGACAGAGAGCTGGGTC 1440
2069 GACCTGAGCTAGAGGGGCTTGGTTCGCGCAATTAATGATTTGACAGAGAGCTGGGTC 2128
1441 CGTGCATATCCACAGAGAGACAGCTGGGCGGCTGTGGCAGGCTGAACACCTGGGAG 1500
2129 CGTGCATATCCACAGAGAGACAGCTGGGCGGCTGTGGCAGGCTGAACACCTGGGAG 2188
1501 CGGCTGAGATCGGCTTTCCTGTGACCCAGACCTGATTCGACACGAGGCAACCCCTTACA 1560
2189 CGGCTGAGATCGGCTTTCCTGTGACCCAGACCTGATTCGACACGAGGCAACCCCTTACA 2248
1561 GGTAGGAGAGAACAGCGGCTGTGCTCCACCTAGAGGCGCTTGTGCGGCGCTCCACTG 1620
2249 GGTAGGAGAGAACAGCGGCTGTGCTCCACCTAGAGGCGCTTGTGCGGCGCTCCACTG 2308
1621 GAAAGTCTCAGTTCCTCTCAGAGCTTCTAGAACATCTGGGCAAGGCTCATGGCTGAT 1680
2309 GAAAGTCTCAGTTCCTCTCAGAGCTTCTAGAACATCTGGGCAAGGCTCATGGCTGAT 2368
1681 AATTTCCCTAGGCTTAAACCCAGCAAGCTTGGGCTGTGCTGTTTATTTTGGTTAA 1740
2369 AATTTCCCTAGGCTTAAACCCAGCAAGCTTGGGCTGTGCTGTTTATTTTGGTTAA 2428
1741 CTTATGAAGAAATGATTAAGAAAGAGAGCTGAGAGAGATTCAGAGATGGAACACACC 1800
2429 CTTATGAAGAAATGATTAAGAAAGAGAGCTGAGAGAGATTCAGAGATGGAACACACC 2488
1801 AGACCCCAAGATCAAAAGCCAAACATGCCCCCTCCAGCAACCCCAAGCCCAAGAC 1860
2489 AGACCCCAAGATCAAAAGCCAAACATGCCCCCTCCAGCAACCCCAAGCCCAAGAC 2548
1861 CATGCTTCTGAATTTCTGAGACACCGTGAAGCTGCTTGTATCTTAACTAGAGAG 1920
2549 CATGCTTCTGAATTTCTGAGACACCGTGAAGCTGCTTGTATCTTAACTAGAGAG 2608
1921 AATACTACCTTCAAGCTTTTGAAGAAATATTTTCTGTTG 1959
2609 AATACTACCTTCAAGCTTTTGAAGAAATATTTTCTGTTG 2647

RESULT 12
AX775949 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775949
DEFINITION Sequence 219 from Patent WO03048202.
ACCESSION AX775949
VERSION AX775949.1 GI:32693667
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Matoudu, A. and Muramatsu, S.
TITLE NF- κ B activating gene
JOURNAL Patent: WO 03048202-A 219 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. 2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755. 2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB11674.1"
/db_xref="GI:32693669"
CDS
1
/translation="MDVYEVASGSMWAGREDIIMKYEKHRAGLPEBDKPKFPSTNN
NVDHIGIVHETELPPLTAREAKQIREISRSKSWMDLGDWERYKSRLLIDRAKGM
PMNIRGPMWSVLTNTEEMKLKNPGRVQIMKEGKRSBHIQRIIDVSGTLKHIFER
DRYTKORELHLIILAYEYNEVEVQCDLSIALFLYLPEBDAFMALVOLASER
HSLOGHSPNGSTVOGLDQDOEHVAVATSOPTMGHIOISGLTLRLMDVYLVEBOLM
PTRLAFYVQOQRLLTKTSRCGPWAFPCRPUDTWARDEDTYUKHLRASMKULTRQGD
IQPKAPKQSSASRPVPASRGKTLCKGRQAPGPFPARFPRLWSASPPRPNST
PCPGAVBEDTYPVGTQVPSPALAQGGQSWRLQWMSMRPLPLDLVEBPFMHY
DFRQSCWRAISQEDQLAPCWQAEHPARVSAFAPASTDSQGTPIFRARDPCAPAT
SGPCLGHLBSSQPPPGF"
ORIGIN
Query Match 92.0%; Score 1807.4; DB 6; Length 2647;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
1 ATGACGAGGTAGAGTGGCGGGCAGTTGTGTGGGCAAGAGCGAGAGCAATCATTTAG 60
755 ATGACGAGGTAGAGTGGCGGGCAGTTGTGTGGGCAAGAGCGAGAGCAATCATTTAG 814
61 AAATACGAAAGGAGACACCGAGCTGGGCTGCCAGAGGCAAGAGGGCTTAAGCTTTTCA 120
815 AAATACGAAAGGAGACACCGAGCTGGGCTGCCAGAGGCAAGAGGGCTTAAGCTTTTCA 874
121 AGCTAACAAACACGTCGATCATTTGGGGATTTGTACATGAGACGAGAGCTGCTCTCTG 180
875 AGCTAACAAACACGTCGATCATTTGGGGATTTGTACATGAGACGAGAGCTGCTCTCTG 934
181 ACTGGCGGGAGGCGCAAGCAATTTGGGGAGATCACCGCAAGAGCAAGTGGGTGAT 240
935 ACTGGCGGGAGGCGCAAGCAATTTGGGGAGATCACCGCAAGAGCAAGTGGGTGAT 994
241 ATGCTGGAGACTGGGAGAAATATCAAAAGACAGCAAAAGCTCATAGATGAGCGTAAAG 300
995 ATGCTGGAGACTGGGAGAAATATCAAAAGACAGCAAAAGCTCATAGATGAGCGTAAAG 1054
301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGATGCTCTCTGAAACATTGAGGAATG 360
1055 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGATGCTCTCTGAAACATTGAGGAATG 1114
361 AAGTTGAAAACCCCGGAAAGATACCAATCATGAAGAGAGGCAAGAGTCACTGAG 420
1115 AAGTTGAAAACCCCGGAAAGATACCAATCATGAAGAGAGGCAAGAGTCACTGAG 1174
421 CACATCCAGCGCATGACCGGGACGTAAAGCGGACATTAAGAGACATATATCTTCAAG 480
1175 CACATCCAGCGCATGACCGGGACGTAAAGCGGACATTAAGAGACATATATCTTCAAG 1234
481 GATCGATACGGAACCAAGCAGCGGAACTACTCCATCTCTCTGATATGAGAGAT 540
1235 GATCGATACGGAACCAAGCAGCGGAACTACTCCATCTCTCTGATATGAGAGAT 1294
541 AACCCGAGGTGGGTACTGACAGGACCTGAGCCACATGCGCGCTTCTCTCTAT 600

```
|||||
Db      1295 AACCCGAGGTGGCTACTGACAGGACCTAGCCACATCCCGCTTGTCTCTCTAT 1354
QY      601 CTTCGAGGAGATGATCTGGGCACTGGTGCAGTGTGGCCAGTGAAGCACTCC 660
Db      1355 CTTCGAGGAGATGATCTGGGCACTGGTGCAGTGTGGCCAGTGAAGCACTCC 1414
QY      661 CTGACGAGATTTTCAAGCCCAATGCGGGACCTCCAGGGGCTCCAAAGCAACAGAG 720
Db      1415 CTGACGAGATTTTCAAGCCCAATGCGGGACCTCCAGGGGCTCCAAAGCAACAGAG 1474
QY      721 CATGTGTAGCCACGTCAACCCAGAGCATGSGGCAATGAGCAAGAAAGATCTATGT 780
Db      1475 CATGTGTAGCCACGTCAACCCAGAGCATGSGGCAATGAGCAAGAAAGATCTATGT 1515
QY      781 GGGCAGTGTCCCGTGAAGCTGCATCCGGAATTTGATGACGGGATCTGCTCGGG 840
Db      1516 -----GATCTCTCGGG 1528
QY      841 CTCACCTGCGCCTGTGGAGAGTGTATCTGTAGAAAGCAAGCGTGTGATGCCGATA 900
Db      1529 CTCACCTGCGCCTGTGGAGAGTGTATCTGTAGAAAGCAAGCGTGTGATGCCGATA 1588
QY      901 ACAAGAAATGCTTTAAAGTTTCAAGACGCTCAAGAAACGTCCAGGTGTGCTCG 960
Db      1589 ACAAGAAATGCTTTAAAGTTTCAAGACGCTCAAGAAACGTCCAGGTGTGCTCG 1648
QY      961 TGGGCAAGTTTGGCAACCGGTTGTTGATACCGGGGCGAGGATGAGACACTGTGCTC 1020
Db      1649 TGGGCAAGTTTGGCAACCGGTTGTTGATACCGGGGCGAGGATGAGACACTGTGCTC 1708
QY      1021 AAGCATCTTAGGCTCTATGAAAGAACTAAAGAAAGAGGAGGAGCTGCCACCCCA 1080
Db      1709 AAGCATCTTAGGCTCTATGAAAGAACTAAAGAAAGAGGAGGAGCTGCCACCCCA 1768
QY      1081 GCCAAACCCGAGCAAGGCTGTGGCATCCAGGCTGTGCCGCTTCACTGTGCGGAG 1140
Db      1769 GCCAAACCCGAGCAAGGCTGTGGCATCCAGGCTGTGCCGCTTCACTGTGCGGAG 1828
QY      1141 ACCCTGTGCAAGGGGGGCAAGGAGGCGCCCTCAAGGCCCAACAGCCGCGTCCGCGGCCC 1200
Db      1829 ACCCTGTGCAAGGGGGGCAAGGAGGCGCCCTCAAGGCCCAACAGCCGCGTCCGCGGCCC 1888
QY      1201 ATTGTGTACGCTTCCCGCAGCGGCACTGTCTTCCACACCTGTGCTGTGAGGCT 1260
Db      1889 ATTGTGTACGCTTCCCGCAGCGGCACTGTCTTCCACACCTGTGCTGTGAGGCT 1948
QY      1261 GTCCGGAGAGACACTTACCTGTGGGCACTCAGGGTGTGCCAGCCGCGCTGTGCTCAG 1320
Db      1949 GTCCGGAGAGACACTTACCTGTGGGCACTCAGGGTGTGCCAGCCGCGCTGTGCTCAG 2008
QY      1321 GAGAGACCTGAGGTTCTGTGAGATTCCTGAGTGAATCCATGCCCCCGCTCCCAAG 1380
Db      2009 GAGAGACCTGAGGTTCTGTGAGATTCCTGAGTGAATCCATGCCCCCGCTCCCAAG 2068
QY      1381 GACCTGAGCTAGAGGGCCCTTGGTTCCGCATTATGATTTAGACAGAGCTGTGAGT 1440
Db      2069 GACCTGAGCTAGAGGGCCCTTGGTTCCGCATTATGATTTAGACAGAGCTGTGAGT 2128
QY      1441 CGTGCCATATCCAGAGAGACAGAGTGGCCCCCTGTGCGAGAGCTGAAACCTTGCGGAG 1500
Db      2129 CGTGCCATATCCAGAGAGACAGAGTGGCCCCCTGTGCGAGAGCTGAAACCTTGCGGAG 2188
QY      1501 CGGGTGAATCGGCTTCCGTCGACCCAGACATGATTCGACCAAGGGGCAACCCCTTCAGA 1560
Db      2189 CGGGTGAATCGGCTTCCGTCGACCCAGACATGATTCGACCAAGGGGCAACCCCTTCAGA 2248
QY      1561 GCTAGGAGCAAGACGCTGTGCTCCACCTCAGGGGCTTGTGCTGTGCGGCTCACTTG 1620
Db      2249 GCTAGGAGCAAGACGCTGTGCTCCACCTCAGGGGCTTGTGCTGTGCGGCTCACTTG 2308
QY      1621 GAAAGTTCTAGTTCCCTCAAGGCTTCTAAGAGCATGTGGGCCAGGGCTCAGTGGCTGAT 1680
|||||
```

```
Db      2309 GAAAGTTCTAGTTCCCTCCAGGCTTCTAAGAGCATGTGGGCCAGGGCTTCAGCTGAGT 2368
QY      1681 AATTTCCCTAGGCTTAACAACCAAGCAAGCTTGGGCTCGTTTATTTTGGTTAA 1740
Db      2369 AATTTCCCTAGGCTTAACAACCAAGCAAGCTTGGGCTCGTTTATTTTGGTTAA 2428
QY      1741 CTTATGAAATGTATTAAAGAAAGTGCAGCTTGAAGAGATTCAGAGATGAAACACC 1800
Db      2429 CTTATGAAATGTATTAAAGAAAGTGCAGCTTGAAGAGATTCAGAGATGAAACACC 2488
QY      1801 AGACCCCAATCAACAAGCAACATGCCCCCGCTCCAGAGACCCCAAGCCCAAGC 1860
Db      2489 AGACCCCAATCAACAAGCAACATGCCCCCGCTCCAGAGACCCCAAGCCCAAGC 2548
QY      1861 CATGTTCTGATTTGACGACACCGTGAAGCTGCTGTTGATCTTAACTCATGAGAG 1920
Db      2549 CATGTTCTGATTTGACGACACCGTGAAGCTGCTGTTGATCTTAACTCATGAGAG 2608
QY      1921 ATAACTACCTTCAAGTTTGAATAAATGTTTCTGTG 1959
Db      2609 ATAACTACCTTCAAGTTTGAATAAATGTTTCTGTG 2647

RESULT 13
AX833552
LOCUS      AX833552      2647 bp      DNA      linear      PAT 15-DEC-2003
DEFINITION      Sequence 676 from Patent EP1347046.
ACCESSION      AX833552
VERSION      AX833552.1 GI:39919687
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homiidae; Homo.

REFERENCE
1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
  Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
  Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
  Masuko, Y.
  Full-length cDNA sequences
  Patent: EP 1347046-A 676 24-SEP-2003;
  Research Association for Biotechnology (JP)

TITILE      JOURNAL
JOURNAL      Location/Qualifiers

FEATURES
source      1..2647
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"

ORIGIN
Query Match      92.0%; Score 1807.4; DB 6; Length 2647;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY      1 ATGAGCTGTGTAGAGGTCCGGGCACTGTGTGGGCAAGAGCGAGAGACATCATTTATG 60
Db      755 ATGAGCTGTGTAGAGGTCCGGGCACTGTGTGGGCAAGAGCGAGAGACATCATTTATG 814
QY      61 AATTAAGAAAAGGAGACACGAGTGGCTGCCAGAGAGAAAGGGGCTTAAGCCTTTTCCA 120
Db      815 AATTAAGAAAAGGAGACACGAGTGGCTGCCAGAGAGAAAGGGGCTTAAGCCTTTTCCA 874
QY      121 AGCTACAAACAACAAGTCGATCTATTGGGGATTTGTACATGAGACGAGAGTGCCTCTCTG 180
Db      875 AGCTACAAACAACAAGTCGATCTATTGGGGATTTGTACATGAGACGAGAGTGCCTCTCTG 934
QY      181 ACTGCCGGGAGGCGAAGCAAAATTCGGCGGGAGATCAAGCGAAAAGCAAGTGCGTGGAT 240
Db      935 ACTGCCGGGAGGCGAAGCAAAATTCGGCGGGAGATCAAGCGAAAAGCAAGTGCGTGGAT 994
QY      241 ATGCTGGGAGACTGGGAGAAATTCAAAAGCAGAGAAAGCTCTAATGAGAGGTCAAG 300
Db      995 ATGCTGGGAGACTGGGAGAAATTCAAAAGCAGAGAAAGCTCTAATGAGAGGTCAAG 1054
|||||
```


QY 301 GGAATGCCCATGAACATCCGCGGCCCCGATGTGCTCCTCTGAACATTGAGGAATG 360
 DB 1055 GGAATGCCCATGAACATCCGCGGCCCCGATGTGCTCCTCTGAACATTGAGGAATG 1114
 QY 361 AAGTTGAAAAACCCCGGAGATATCAATCATGAAAGAAAGGCAAGATCATCTGAG 420
 DB 1115 AAGTTGAAAAACCCCGGAGATATCAATCATGAAAGAAAGGCAAGATCATCTGAG 1174
 QY 421 CACATTCAGCGCATGCAACCGGAGCTGTAAGCGGAGCATTAAGAAAGATATTTCTGAG 480
 DB 1175 CACATTCAGCGCATGCAACCGGAGCTGTAAGCGGAGCATTAAGAAAGATATTTCTGAG 1234
 QY 481 GATGATATGGAACCAAGAGCGGAGAACTACTCCATCTCTCTGCAATATGAGAGTAT 540
 DB 1235 GATGATATGGAACCAAGAGCGGAGAACTACTCCATCTCTCTGCAATATGAGAGTAT 1294
 QY 541 AACCGGAGGTGGGCTACTGACAGGAGCTGAGCAATGCGCGCTTGTCTCTCTAT 600
 DB 1295 AACCGGAGGTGGGCTACTGACAGGAGCTGAGCAATGCGCGCTTGTCTCTCTAT 1354
 QY 601 CTTCCTGAGAGAGATGCAATTTCTGAGCACTGTGTGAGCTGTGCGCAGTGAAGGCACTCC 660
 DB 1355 CTTCCTGAGAGAGATGCAATTTCTGAGCACTGTGTGAGCTGTGCGCAGTGAAGGCACTCC 1414
 QY 661 CTGCAAGGATTTCAAGCCCAAAATGCGCGGAGCTGCAAGGAGCTCCAAAGCAACAGAG 720
 DB 1415 CTGCAAGGATTTCAAGCCCAAAATGCGCGGAGCTGCAAGGAGCTCCAAAGCAACAGAG 1474
 QY 721 CATGTGTAGCCACGTCACAAACCAAGACCATGTGGGATCAGGAAACAAGAAATCTATGT 780
 DB 1475 CATGTGTAGCCACGTCACAAACCAAGACCATGTGGGATCAGGAAACAAGAAATCTATGT 1515
 QY 781 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGAGATATGATGAGCGGATCTCTCTCGGG 840
 DB 1516 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGAGATATGATGAGCGGATCTCTCTCGGG 1528
 QY 841 CTACCCCTGCGCTGTGGGACGTTATCTGTATGAAAGGGAACAGGCGTTGATGCGGAAT 900
 DB 1529 CTACCCCTGCGCTGTGGGACGTTATCTGTATGAAAGGGAACAGGCGTTGATGCGGAAT 1588
 QY 901 ACAAGAAATGCGCTTTAAGGTTTCAAGAGAAAGCCCTTCAAGAAAGCTGAGTGTGGCCG 960
 DB 1589 ACAAGAAATGCGCTTTAAGGTTTCAAGAGAAAGCCCTTCAAGAAAGCTGAGTGTGGCCG 1648
 QY 961 TGGGCAAGTTTGTGCAACCGGTTCTGTATACCTGTGGCCAGGAGATGAGCACTGTGCTC 1020
 DB 1649 TGGGCAAGTTTGTGCAACCGGTTCTGTATACCTGTGGCCAGGAGATGAGCACTGTGCTC 1708
 QY 1021 AAGCATTTTAAAGGCTCTATGAAAGAACTTAACAAGAAAGAGGGGAGCTTCCACCCCA 1080
 DB 1709 AAGCATTTTAAAGGCTCTATGAAAGAACTTAACAAGAAAGAGGGGAGCTTCCACCCCA 1768
 QY 1081 GCCAAACCCGAGCAAGAGGCTGTGCGCATCAGAGCCCTGATCGGAGCTTCACTGTGCGGGAAG 1140
 DB 1769 GCCAAACCCGAGCAAGAGGCTGTGCGCATCAGAGCCCTGATCGGAGCTTCACTGTGCGGGAAG 1828
 QY 1141 ACCCTCTGCAAGGGGAGCAAGGAGCCCTTCAAGGCGCAACAGCCCGGTTTCCCGGGCCC 1200
 DB 1829 ACCCTCTGCAAGGGGAGCAAGGAGCCCTTCAAGGCGCAACAGCCCGGTTTCCCGGGCCC 1888
 QY 1201 ATTGTGTAGCTTCCCGGCAAGGGGACCTGTTCTTCAACACCTGTGTCTGTGTGGGCT 1260
 DB 1889 ATTGTGTAGCTTCCCGGCAAGGGGACCTGTTCTTCAACACCTGTGTCTGTGTGGGCT 1948
 QY 1261 GTCCGGGAAAGACACCTTACCTGTGGGACCTCAAGGAGTGTCCAGCCCGGCTGTGCTAG 1320
 DB 1949 GTCCGGGAAAGACACCTTACCTGTGGGACCTCAAGGAGTGTCCAGCCCGGCTGTGCTAG 2008
 QY 1321 GGAGAGCTCAGAGGTTCTGAGATTTCTGAGATGAGAACTCAATGCCCCGCTCCCAAG 1380
 DB 2009 GGAGAGCTCAGAGGTTCTGAGATTTCTGAGATGAGAACTCAATGCCCCGCTCCCAAG 2068

QY 1381 GACCTGAGAGTGAAGGGCCCTTGTGTTCCGCTATTAATGATTTGACAGAGAGCTGTGGTCC 1440
 DB 2069 GACCTGAGAGTGAAGGGCCCTTGTGTTCCGCTATTAATGATTTGACAGAGAGCTGTGGTCC 2128
 QY 1441 CGTGCATATTCACAGAGAGACCAAGCTGTGCCCCCTGTGACAGCTGAAACACCTGTGGAG 1500
 DB 2129 CGTGCATATTCACAGAGAGACCAAGCTGTGCCCCCTGTGACAGCTGAAACACCTGTGGAG 2188
 QY 1501 CGGCTGAGATCGGCTTTGCTGCTGACACCCAGCACTGATTTCCGACAGGGCACCCCTTCA 1560
 DB 2189 CGGCTGAGATCGGCTTTGCTGCTGACACCCAGCACTGATTTCCGACAGGGCACCCCTTCA 2248
 QY 1561 GCTAGGAGACCAAGCGGCTGTGCTCCACCTCAAGGGCCCTGTGCTGTGCGGCTCCACTTG 1620
 DB 2249 GCTAGGAGACCAAGCGGCTGTGCTCCACCTCAAGGGCCCTGTGCTGTGCGGCTCCACTTG 2308
 QY 1621 GAAAGTTCTCAATTCCTTCCACAGGCTTCAAGAGCATCTGTGGCCAGGCTCATGTGAT 1680
 DB 2309 GAAAGTTCTCAATTCCTTCCACAGGCTTCAAGAGCATCTGTGGCCAGGCTCATGTGAT 2368
 QY 1681 AATTTCCCTAGGCTTAAACCAACCAAGCAAGCTTCCGCTCTGTTTATTTTGTGTTAA 1740
 DB 2369 AATTTCCCTAGGCTTAAACCAACCAAGCAAGCTTCCGCTCTGTTTATTTTGTGTTAA 2428
 QY 1741 CTTATGAAATGATTTAAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800
 DB 2429 CTTATGAAATGATTTAAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2488
 QY 1801 AGACCCAGATCAACAAGCCCAACATGCCCCCTCCAGACACCCCAAGCCCAAGCAGAC 1860
 DB 2489 AGACCCAGATCAACAAGCCCAACATGCCCCCTCCAGACACCCCAAGCCCAAGCAGAC 2548
 QY 1861 CATGCTTGAATTTCTGACAGACACCTGAGCTGCTTGTATCTTAACTATGAGAG 1920
 DB 2549 CATGCTTGAATTTCTGACAGACACCTGAGCTGCTTGTATCTTAACTATGAGAG 2608
 QY 1921 ATATCTACCTTCAAGCTTTTGAATTAATGTTCTCTGTTG 1959
 DB 2609 ATATCTACCTTCAAGCTTTTGAATTAATGTTCTCTGTTG 2647

RESULT 14
 AK095385
 LOCUS 2647 bp mRNA linear PRI 30-JUN-2004
 DEFINITION Homo sapiens cDNA FLJ38066 fis. clone CTONG2015330, moderately
 similar to H.sapiens mRNA for lre oncogene.
 ACCESSION AK095385
 VERSION AK095385.1 GI:21754631
 KEYWORDS oligo capting, fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Wakematsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
 Sekine, M., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
 Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
 Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
 Tanikawa, M., Yamazaki, M., Niimiyi, K., Ishibashi, T., Yamashita, H.,
 Hironaka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Matanabe, S.,
 Yoisida, M., Horiuchi, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
 Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Muesshino, K., Yunkai, H., Oshima, A.,
 Saeki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
 Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satch, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Makebe, H., Hishigaki, H., Matanabe, T.,

Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumaishi, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiroe, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togsashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuo, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL PUBMED 14702039

TITLE 2

REFERENCE 2

AUTHORS Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hiroe, M., Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Mateno, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Watsushima, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuo, Y., Nagai, K., and Isogai, T.

TITLE 3 (bases 1 to 2647)

JOURNAL Unpublished

REFERENCE Isogai, T. and Yamamoto, J.

AUTHORS Direct Submission

TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7 Kazuura-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.

FEATURES Location/Qualifiers

source 1. 2647

organism="Homo sapiens"

molecule="mRNA"

db_xref="taxon:9606"

clone="CTONG2015330"

tissue="tongue, tumor tissue"

clone_lib="CTONG2"

note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 92.0% Score 1807.4; DB 8; Length 2647;

Best Local Similarity 96.3% Pred. No. 0;

Matches 1887; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

1 ATGACGCTGTAGAGTCCGGCGAGCTTGGTGGCAACAAGCGAGCATCATTTATG 60

DB ATGACGCTGTAGAGTCCGGCGAGCTTGGTGGCAACAAGCGAGCATCATTTATG 814

61 AAATACGAAAGGAGACACGAGCTGGCTCCAGAGGACAAAGGGCTTAAGCTTTTGA 120

DB AAATACGAAAGGAGACACGAGCTGGCTCCAGAGGACAAAGGGCTTAAGCTTTTGA 874

815 AAATACGAAAGGAGACACGAGCTGGCTCCAGAGGACAAAGGGCTTAAGCTTTTGA 874

121 AGCTACAAACAACGCTGATCATTTTGGGATTTGTACATGACGAGCTGCTCTCTG 180

DB AGCTACAAACAACGCTGATCATTTTGGGATTTGTACATGACGAGCTGCTCTCTG 934

875 AGCTACAAACAACGCTGATCATTTTGGGATTTGTACATGACGAGCTGCTCTCTG 934

181 ACTGCGGAGGAGGAGCAAAATTCGGCGGAGATCGCGAAGAGCAAGTGGGTGAT 240

DB ACTGCGGAGGAGGAGCAAAATTCGGCGGAGATCGCGAAGAGCAAGTGGGTGAT 994

935 ACTGCGGAGGAGGAGCAAAATTCGGCGGAGATCGCGAAGAGCAAGTGGGTGAT 994

241 ATGCTGGAGACTGGAGAAATACAAAGACGAGAAAGCTCATAGATGAGGCTTAAG 300

DB 995 ATGCTGGAGACTGGAGAAATACAAAGACGAGAAAGCTCATAGATGAGGCTTAAG 1054

QY 301 GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTGCTCTGAACATTTGAGAAATG 360

DB 1055 GGAATGCCATGAACATCCGGGGCCGATGTGGTCACTGCTCTGAACATTTGAGAAATG 1114

QY 361 AAGTTAAAAAACCCTGGAATATCAAGATCATAGAGAGAGAGGCAAGAGTCACTTGA 420

DB 1115 AAGTTAAAAAACCCTGGAATATCAAGATCATAGAGAGAGAGGCAAGAGTCACTTGA 1174

DB 421 CACATCCAGCGCATCGACCGGGACGTAACGGGACATTAAGAGACATATATTTCCAG 480

QY 1175 CACATCCAGCGCATCGACCGGGACGTAACGGGACATTAAGAGACATATATTTCCAG 1234

DB 481 GATCATGAGGAACCAAGACGCGGGAATCTACATCATCTCTGTCATATGAGAGTAT 540

QY 1235 GATCATGAGGAACCAAGACGCGGGAATCTACATCATCTCTGTCATATGAGAGTAT 1294

DB 541 AACCCGAGGTGGCTATGTCAGAGGACCTGAGCCATCGCCCTTGTCTCTCTAT 600

QY 1295 AACCCGAGGTGGCTATGTCAGAGGACCTGAGCCATCGCCCTTGTCTCTCTAT 1354

DB 601 CTTCTGAGAGGATGATCTTGGGCACTGTGTCACTGTGCTGGCCAGTGAAGGCACTCC 660

QY 1355 CTTCTGAGAGGATGATCTTGGGCACTGTGTCACTGTGCTGGCCAGTGAAGGCACTCC 1414

DB 661 CTGACGAGATTTACAGCCCAATGCGGGGACCGTCCAGGGGCTCCAAAGCAACAGAG 720

QY 1415 CTGACGAGATTTACAGCCCAATGCGGGGACCGTCCAGGGGCTCCAAAGCAACAGAG 1474

DB 721 CATGTGTAGCAAGTCAACACCAAGACCATGAGGATCAGAGCAAAAGATCTATGT 780

QY 1475 CATGTGTAGCAAGTCAACACCAAGACCATGAGGATCAGAGCAAAAGATCTATGT 1515

DB 781 GGGCAATGTTCCCGTTAGGCTGCTCATCCGATTTGATTAAGGATCTCTCTGGG 840

QY 1516 GGGCAATGTTCCCGTTAGGCTGCTCATCCGATTTGATTAAGGATCTCTCTGGG 1528

DB 841 CTGACCTGGGCTGTGGGACGTATCTGTGTAAGAGCAACAGGCTTATGATGCGGAT 900

QY 1529 CTGACCTGGGCTGTGGGACGTATCTGTGTAAGAGCAACAGGCTTATGATGCGGAT 1588

DB 901 ACAAGATGCGCTTTAAGTTTCAGAGAGCGCTTCAGAGACGTCAGAGTGGCCCG 960

QY 1589 ACAAGATGCGCTTTAAGTTTCAGAGAGCGCTTCAGAGACGTCAGAGTGGCCCG 1648

DB 961 TGGGCACTGTTTTCAGACCGGCTGTTGATCTTGGGCCAGGATGAGACATGTGCTC 1020

QY 1649 TGGGCACTGTTTTCAGACCGGCTGTTGATCTTGGGCCAGGATGAGACATGTGCTC 1708

DB 1021 AAGCATCTTAGGGCTCTATGAGAACTAACAAAGAAAGAGGGAGACCTGCCACCCCA 1080

QY 1709 AAGCATCTTAGGGCTCTATGAGAACTAACAAAGAAAGAGGGAGACCTGCCACCCCA 1768

DB 1081 GCCAAACCGAGAGAGGCTGTGCGATCCAGGCTGTGCGGCTTCACTGTGCGGAG 1140

QY 1769 GCCAAACCGAGAGAGGCTGTGCGATCCAGGCTGTGCGGCTTCACTGTGCGGAG 1828

DB 1141 ACCCTCTGCAAGGGGACAGGACGCGCTTCAGAGCCCAACAGCCGGTTCGCGGCC 1200

QY 1829 ACCCTCTGCAAGGGGACAGGACGCGCTTCAGAGCCCAACAGCCGGTTCGCGGCC 1888

DB 1201 ATTGTGACGCTTCCCGGACAGGGGACGCTGTTTCCACACCTGTCTGTGAGGAGCT 1260

QY 1889 ATTGTGACGCTTCCCGGACAGGGGACGCTGTTTCCACACCTGTCTGTGAGGAGCT 1348

DB 1261 GTCGAGGAGACACTTACCTGTGTGCACTCAAGGTGTGCCACGCGGCTGTGCTCAG 1320

QY 1949 GTCGAGGAGACACTTACCTGTGTGCACTCAAGGTGTGCCACGCGGCTGTGCTCAG 2008

DB 1321 GAGGACCTCAGGGTTCGTGAGATTTCTGCACTGGAATCTCATGCTCCGCTCCAAAG 1380

QY 2009 GAGGACCTCAGGGTTCGTGAGATTTCTGCACTGGAATCTCATGCTCCGCTCCAAAG 2068

DB

QY	1381	GACCTGAGCGTAAAGGGGCGCTTGCTCGGCATTATGATTTTCAACAAGCGCTGGGCTC	1441.0
Db	2069	GACCTGAGCGTAAAGGGGCGCTTGCTCGGCATTATGATTTTCAACAAGCGCTGGGCTC	2128.9
QY	1441	CGTGGCATATCCCAAGAGGAGCAAGCTGGGCGCCCTGCTGGCAGGCTGAACAACCTGCGAG	1500.0
Db	2129	CGTGGCATATCCCAAGAGGAGCAAGCTGGGCGCCCTGCTGGCAGGCTGAACAACCTGCGAG	2188.9
QY	1501	CGGGTGAGATGGGCTTTTGGCTGGCACCCAGCACTGATTTCCGACCAAGGGCAACCCCTTCAGA	1560.0
Db	2189	CGGGTGAGATGGGCTTTTGGCTGGCACCCAGCACTGATTTCCGACCAAGGGCAACCCCTTCAGA	2248.9
QY	1561	GCTAGGGAGCAACAGCGGTGTGCTCCGACTCAAGGGGCGTTGCTGTGGGGCTCCACTTG	1620.0
Db	2249	GCTAGGGAGCAACAGCGGTGTGCTCCGACTCCAGGGGCGTTGCTGTGGGGCTCCACTTG	2308.9
QY	1621	GAAGGTTCTCAGTTCCTCTCAAGGCTTCTAGAAGCATCTGGGCGAAGGCTCACTAGCTGGAT	1680.0
Db	2309	GAAGGTTCTCAGTTCCTCTCAAGGCTTCTAGAAGCATCTGGGCGAAGGCTCACTAGCTGGAT	2368.9
QY	1681	AATTTTCCCTAGGCTTAACAAACCCAAAGCAAGCTTGCGCTCTCGTTTATTTTGGTTAA	1740.0
Db	2369	AATTTTCCCTAGGCTTAACAAACCCAAAGCAAGCTTGCGCTCTCGTTTATTTTGGTTAA	2428.9
QY	1741	CTTATGAAAAATGTATTTAAGAAAGGTGAGCTCGACAGAGATTCAGAGATGGAACACACC	1800.0
Db	2429	CTTATGAAAAATGTATTTAAGAAAGGTGAGCTCGAGAGATTCAGAGATGGAACACACC	2488.9
QY	1801	AGACCCCGAGTCAACAAAGCCAAACCATGCGCCAGCCCTCCAGACACCCCGACGCCACGAC	1860.0
Db	2489	AGACCCCGAGTCAACAAAGCCAAACCATGCGCCAGCCCTCCAGACACCCCGACGCCACGAC	2548.9
QY	1861	CATCGTTCTGAATTTCTGACGACACCGTAGGCTGCTTTGACTTTAACTCATGGAAG	1920.0
Db	2549	CATCGTTCTGAATTTCTGACGACACCGTAGGCTGCTTTGACTTTAACTCATGGAAG	2608.9
QY	1921	ATAACTACTTCACGTTTGTGAATTAATGTTTCTCGTTG	1959
Db	2609	ATAACTACTTCAGTTTGTGAATTAATGTTTCTCGTTG	2647

[illegible]

CONSRM	Schneerch, A., Schein, J.E., Jones, S.J. and Marz, M.A.
TITLE	Mammalian Gene Collection Program Team
JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE	12477932
AUTHORS	2 (bases 1 to 2111)
CONSRM	NIH MGC Project
TITLE	Direct Submission
JOURNAL	Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcgaps-remail.nih.gov Issue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcsc.bc.ca Martin Hirst, Thomas Zeng, Ryan Morin, Michelle Moka, Johnson Pang, Diana Mah, Jing Wang, Kiehl Fischer, Eric Chuah, Allen Delaney, Rob Kirkpatrick, Agnes Baross, Sarah Barber, Mabel Brown-John, Steve S. Chand, William Chow, Ryan Babkaiff, Dave Wong, Corey Marcuo, Jaclyn Beland, Susan Gibson, Luis delRio, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Mike R. Mayo, Josh Moran, Diana Palmquist, JR Santos, Duane Smalish, Jeff Stott, Miranda Teali, George Yang, Jacques Schein, Asim Siddiqui, Steven Jones, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 119 Row: 1 Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.

location/Qualifiers

1. 2111

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:64921 IMAGE:5744726"
/issue type="Brain, adult medulla"
/clone_id="NIH MGC_119"
/lab host="DH10B"
/note="Vector: pCMV-SPORT6"

1. 2111

/gene="LOC40452"
/db_xref="GeneID:440452"
85_1125

/gene="LOC40452"
/codon start=1
/product="hypothetical protein LOC40452"
/protein_id="AAHS8890.1"
/db_xref="GI:37590618"
/db_xref="GeneID:440452"
/translation="MDVVVEYQSWMAOERDITMKYEKGKRAGLPDKGPGSGVYN
NIDHFQMCHTELPPTARVKQIRRBISRSKSKFWKMLGSWDTTKSRKLIDAYOSI
PMNIRQEMSLVLINIEIKLNKGRYOIMKEKRSSEHIQQMDLVSGTTRRIIFER
DRGTQRRELVTILALIEEYNPEGVCRDIASHAFLVLPEDBAFWALVOLASRER
HSIQGSPNGVGIGLDOOEHVATSSLPKTMMODKKDLCCGCSGISCLITLDGG
ISGLTLRLMDVYLIVGEQAFLMITSPSPLRFSSASRRRGVARGHVFQGLSLMPGG
MTMLCSIIIGPL"

ORIGIN

Query Match 91.6%; Score 1799.4; DB 8; Length 2111;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1900; Conservative 0; Mismatches 61; Indels 44; Gaps 2;

Oy 1 ATGAGCGTGTTAAGTCGCGGCCACTTGTGGCCACAAGGACAGCATCATTATG 60
|||||
db 85 ATGACGCTGTAAAGCTCGTGGTAGTTGATGGCACACAGAGCAAGCATCATTTATG 144

QY 61 AAATACGAAAAAGGACACCGAGCTGGCTCCAGAGGACAAAGGGGCTTAAGCTTTTCGA 120
DB 145 AAATACGAAAAAGGACACCGAGCTGGCTCCAGAGGACAAAGGGGCTTAAGCTTTTCGA 204
QY 121 AGCTACAAACAACAACGTCGATTCATTTGGGGATTGTATATAGAACGGAGCTGCCCTCTCG 180
DB 205 AGCTACAAACAACAACGTCGATTCATTTGGGGATTGTATATAGAACGGAGCTGCCCTCTCG 264
QY 181 ACTGCGGGGAGGCGAAGCAAAATTCGCGGGGAGATCAGCCGAAAGCAAGCTGGTGTAT 240
DB 265 ACTGCGGGGAGGCGAAGCAAAATTCGCGGGGAGATCAGCCGAAAGCAAGCTGGTGTAA 324
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAAGACGAAAGCTCATATATGTCAGCGTACAG 300
DB 325 ATGCTGGGAGAAATGGGACACTTACAAAACAGCAGAAAGCTCATATATGTCAGCGTACAG 384
QY 301 GGAATGCCATGAAACATCCGGGGCCCGATGTGTGCTCTCTCTGAACTTTAGGAAATG 360
DB 385 GGAATGCCATGAAACATCCGGGGCCCGATGTGTGCTCTCTCTGAACTTTAGGAAATC 444
QY 361 AAGTTGAAAAACCCGGAGATACCAAGATCATGAAAGGAGGCAAGAGTCACTGAG 420
DB 445 AAGTTGAAAAACCCGGAGATACCAAGATCATGAAAGGAGGCAAGAGTCACTGAA 504
QY 421 CACATCCAGGCAATCGACCGGAGAGTAAAGCGGAGCATTTAAGAAACATATATTTCTCAG 480
DB 505 CACATCCAGGCAATCGACCGGAGAGTAAAGCGGAGCATTTAAGAAAGCATATATTTCTCAG 564
QY 481 GATCGATACGGAACCAAGAGCGGGAATCTACATCTCTCTGGATATGAGAGATAT 540
DB 565 GATCGATACGGAACCAAGAGCGGGAATCTACATCTCTCTGGATATGAGAGATAT 624
QY 541 AACCCGGAGGTGGGCTACTGCAAGGACCTGAGCACAATCGCCGCTTTGTTCTCTCTAT 600
DB 625 AACCCGGAGGTGGGCTACTGCAAGGACCTGAGCACAATCGCCGCTTTGTTCTCTCTAT 684
QY 601 CTTCCTGAGGAGGATGATTTCTGGGCACTGTGTGCACTGTGTGCTGTGTGAGGCACTCC 660
DB 685 CTTCCTGAGGAGGATGATTTCTGGGCACTGTGTGCACTGTGTGCTGTGTGAGGCACTCC 744
QY 661 CTGCAAGGATTTTCAACAGCCCAATTTGGGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 720
DB 745 CTGCAAGGATTTTCAACAGCCCAATTTGGGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 804
QY 721 CATGTGTAGCCAGCTGCAACCAACCAAGACCATGTGGGATCAGAGCAAGAAAGATCTATGT 780
DB 805 CATGTGTAGCCAGCTGCAACCAACCAAGACCATGTGGGATCAGAGCAAGAAAGATCTATGT 864
QY 781 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGATATTTGATTTGAACGGGATCTCTCTCGGG 840
DB 865 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGATATTTGATTTGATTTGATTTGATTTGAT 924
QY 841 CTCAACCTGTGCTGTGGGAGCTGTATCTGTGTAGAGGCGAAACAGGGGTTGATGCGGATA 900
DB 925 CTCAACCTGTGCTGTGGGAGCTGTATCTGTGTAGAGGCGAAACAGGGGTTGATGCGGATA 984
QY 901 ACAAGATCGCTTTTAAAGTTTCAAGAGAGCGCTTCAAGAAAGCTCAGAGTGTGGCCG 960
DB 985 ACAAGATCGCTTTTAAAGTTTCAAGAGAGCGCTTCAAGAAAGCTCAGAGTGTGGCCG 1043
QY 961 TGGGCAAGTTTTCGAACCGGTTCTGTTGATACCTGGGCGAGGATGAGACATGTGTCTC 1020
DB 1044 TGGGCAAGTTTTCGAACCGGTTCTGTTGATACCTGGGCGAGGATGAGACATGTGTCTC 1103
QY 1021 AAGCATTTTGGGCTCTTATGAAAGAACTTAAACAAGAAAGCAAGGGGACCTGCCACCCCA 1080
DB 1104 AAGCATTTTGGGCTCTTATGAAAGAACTTAAACAAGAAAGCAAGGGGACCTGCCACCCCA 1163
QY 1081 GCCAAACCCGAGCAAGGGGTGTGGGCAATCCAGGCTGTGGCGGCTTCAAGTGGGAGAG 1140
DB 1164 GCCAAACCCGAGCAAGGGGTGTGGGCAATCCAGGCTGTGGCGGCTTCAAGTGGGAGAG 1223
QY 1141 ACCCTCTGCAAGGGGAGCAAGGCAAGGCCCTTCAGGGCCACCAAGCCCGGTTCCGCGGCC 1200

DB 1224 ACCCTCTGCAAGGGGAGCAAGGCAAGGCCCTTCAGGGCCACCAAGCCCGGTTCCGCGGCC 1283
QY 1201 ATTTGGTCAAGCTTCCCGGACACGGGACACTGTTCTTTCACACACCTGTCTGTGGGGCT 1260
DB 1284 ATTTGGTCAAGCTTCCCGGACACGGGACACTGTTCTTTCACACACCTGTCTGTGGGGCT 1343
QY 1261 GTCCGGGAAAGACACTACCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGGCTCAG 1320
DB 1344 GTCCGGGAAAGACACTACCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGGCTCAG 1403
QY 1321 GAGAGACTCAGGGTCTCTGAGAAATCTGTCAATGTGAATCTGCCCGGCTCCCAAG 1380
DB 1404 GAGAGACTCAGGGTCTCTGAGAAATCTGTCAATGTGAATCTGCCCGGCTCCCAAG 1463
QY 1381 GACCTGAGAGTAAAGGACCTTGGTTCCGCACTTATGATTTTCAGACAGAGCTGTGGGTC 1440
DB 1464 GACCTGAGAGTAAAGGACCTTGGTTCCGCACTTATGATTTTCAGACAGAGCTGTGGGTC 1523
QY 1441 CGTGCAATATCCAGAGAGGACACAGCTGGCCCTGTGTGGAGCTGAAACCTGTGGAG 1500
DB 1524 CGTGCAATATCCAGAGAGGACACAGCTGGCCCACTGTGTGGAGCTGAAACCTGTGGAG 1583
QY 1501 CGGCTGAGATCGGCTTTCTGTGCA----- 1524
DB 1584 CGGCTGAGATCGGCTTTCTGCACTGTGACCAACAGTGGGCAATGACTTCCCGGCCCTG 1643
QY 1525 -----CCAGACACTGATTTCCAGACGAGGACACCCCTTCAGAGCTAGGAGCAAGAGCC 1577
DB 1644 CAGTGGGCCCAAGCACTGATTTCCGACCAAGGGACACCCCTTCAGAGCTAGGAGCAAGAG 1703
QY 1578 GTGTGCTCCCACTCAAGGAGCTTGTGCTGTGGGCTCTCACTTGGAAAGTTCTCAGTTCCC 1637
DB 1704 GTGTGCTCCCACTCAAGGAGCTTGTGCTGTGGGCTCTCACTTGGAAAGTTCTCAGTTCCC 1763
QY 1638 TCCAGGCTTCTAAGACATCTGGGCCAGGGCTCATAGGCTGATTAATTTCCCTAGGCTTAA 1697
DB 1764 TCCAGGCTTCTAAGACATCTGGGCCAGGGCTCATAGGCTGATTAATTTCCCTAGGCTTAA 1823
QY 1698 CAACCAAGCAAGCTTCCGCTCTGCTTTTATTTTGTGTTAACTTATGAAAGTATTA 1757
DB 1824 CAACCAAGCAAGCTTCCGCTCTGCTTTTATTTTGTGTTAACTTATGAAAGTATTA 1883
QY 1758 AGAAGAGTGCAGCTCGAAGAGATTCAGAGATGAAACAACAGACCCCAAGTCAACAA 1817
DB 1884 AGAAGAGTGCAGCTCGAAGAGATTCAGAGATGAAACAACAGACCCCAAGTCAACAA 1943
QY 1818 GCCAACATGCCCCAGCCCTCCAGACACCCAGGCCACAGACATGTTCTGAATTTCTG 1877
DB 1944 GCCAACATGCCCCAGCCCTCCAGACACCCAGGCCACAGACATGTTCTGAATTTCTG 2003
QY 1878 AGCAGACCGTGAAGCTGCTTTGTACTTTTAACTCAATGGAAGGATTACTTCAAGTT 1937
DB 2004 AGCAGACCGTGAAGCTGCTTTGTACTTTTAACTCAATGGAAGGATTAACCACTTCAAGTT 2063
QY 1938 TTGAATTAATGTTTCTGTGAAA 1962
DB 2064 TTGAATTAATGTTTCTGTGAAA 2088

Search completed: April 5, 2006, 19:03:48
Job time : 9950 secs